

Please send to Applicants
Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 15:12:48 ; Search time 590 Seconds
(without alignments)
15500.940 Million cell updates/sec

Title: US-10-030-850-1
Perfect score: 5145
Sequence: 1 gattcgagcagcagcgcgc.....ggatcaaatcaataaaaa 5145

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/1.COMB.seq: *
2: /cgn2_6/prodata/1/ina/5.COMB.seq: *
3: /cgn2_6/prodata/1/ina/6.COMB.seq: *
4: /cgn2_6/prodata/1/ina/6.COMB.seq: *
5: /cgn2_6/prodata/1/ina/6.COMB.seq: *
6: /cgn2_6/prodata/1/ina/6.COMB.seq: *
7: /cgn2_6/prodata/1/ina/6.COMB.seq: *
8: /cgn2_6/prodata/1/ina/6.COMB.seq: *
9: /cgn2_6/prodata/1/ina/6.COMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	732.8	14.2	767	3	US-09-270-767-596 Sequence 596, App
2	732.8	14.2	767	3	US-09-270-767-15878 Sequence 15878, A
3	661	12.8	1420	3	US-09-270-767-6018 Sequence 6018, A
4	661	12.8	1420	3	US-09-270-767-21300 Sequence 21300, A
5	452.4	8.8	466	3	US-09-270-767-14266 Sequence 14266, A
6	149.8	2.9	3153	3	US-10-104-047-327 Sequence 327, App
7	149.4	2.9	154	3	US-09-270-767-30401 Sequence 30401, A
8	141.8	2.3	2524	3	US-10-104-047-517 Sequence 517, App
9	116.6	2.3	1908	3	US-10-104-047-113 Sequence 113, App
10	116.6	2.3	6445	3	US-09-620-312D-324 Sequence 324, App
11	116.6	2.3	6487	3	US-09-620-312D-323 Sequence 323, App
12	47.2	0.9	7218	2	US-08-232-463-14 Sequence 14, App1
13	43.8	0.9	7218	2	US-08-232-463-14 Sequence 14, App1
14	42.8	0.8	3472	2	US-08-232-463-14 Sequence 14, App1
15	41.6	0.8	3684	3	US-09-561-077C-25 Sequence 25, App1
16	41.6	0.8	3684	3	US-09-561-077C-25 Sequence 25, App1
17	41.6	0.8	7113	3	US-09-561-077C-26 Sequence 26, App1
18	41.6	0.8	7113	3	US-09-561-077C-26 Sequence 26, App1
19	41.2	0.8	1713	3	US-09-252-991A-10767 Sequence 10767, A
20	41.2	0.8	2490	3	US-09-252-991A-10579 Sequence 10579, A
21	41.2	0.8	3384	3	US-09-252-991A-10278 Sequence 10278, A
22	41.2	0.8	19124	2	US-08-487-826B-13 Sequence 13, App1
23	41.2	0.8	154746	3	US-09-827-688-8 Sequence 8, App1
24	40.2	0.8	468	3	US-09-252-991A-2152 Sequence 2152, App

C	25	40.2	0.8	777	3	US-09-252-991A-1732 Sequence 1732, App
C	26	40.2	0.8	930	3	US-09-252-991A-1997 Sequence 1997, App
C	27	40	0.8	1290	3	US-09-902-540-5006 Sequence 5006, App
C	28	40	0.8	12950	3	US-09-902-540-1036 Sequence 1036, App
C	29	38.8	0.8	505	3	US-09-621-976-15639 Sequence 15639, A
C	30	38.6	0.8	2358	3	US-09-252-991A-16218 Sequence 16218, A
C	31	38.6	0.8	2358	3	US-09-252-991A-16218 Sequence 16218, A
C	32	38.4	0.7	1026	3	US-09-270-767-12503 Sequence 12503, A
C	33	38.4	0.7	3395	3	US-09-902-540-3658 Sequence 3658, App
C	34	38.4	0.7	23847	3	US-09-902-540-1177 Sequence 1177, App
C	35	38.2	0.7	390416	3	US-09-949-016-16923 Sequence 16923, A
C	36	37.8	0.7	411	3	US-09-252-991A-16270 Sequence 16270, A
C	37	37.6	0.7	996	2	US-08-924-758-21 Sequence 21, App1
C	38	37.6	0.7	996	2	US-08-924-758-21 Sequence 21, App1
C	39	37.6	0.7	2754	3	US-09-489-039A-5411 Sequence 5411, App
C	40	37.4	0.7	1652	3	US-08-726-214-17 Sequence 17, App1
C	41	37.4	0.7	3639	3	US-09-902-540-5646 Sequence 5646, App
C	42	37.4	0.7	3978	3	US-08-726-214-1 Sequence 1, App1
C	43	37.4	0.7	51354	3	US-09-902-540-1270 Sequence 1270, App
C	44	37.2	0.7	50000	3	US-09-662-254B-26 Sequence 26, App1
C	45	37.2	0.7	160759	3	US-09-949-016-16514 Sequence 16514, A

ALIGNMENTS

QY	2639	CCTGATCAGCTTACACAGTATCTGCGCATGGAAGAGCTTACTACAGAGCACTTAA	2639	US-09-270-767-596	1	US-09-270-767-596	14.2%; Score 732.8; DB 3; Length 767;
DB	1	CCTGATCAGCTTACACAGTATCTGCGCATGGAAGAGCTTACTACAGAGCACTTAA	60	Sequence 596, Application US/09270767	1	Patent No. 6703491	Best Local Similarity 98.7%; Pred. No. 8e-208; Matches 760; Conservative 0; Mismatches 7; Indels 3; Gaps 2;
QY	2699	ATGCGTTCGAGCATGTTCCGAGATGTTCAATGATACACACCAAGGTACTACT	2758	GENERAL INFORMATION:	1	APPLICANT: Homburger et al.	
DB	61	ATGCGTTCGAGCATGTTCCGAGATGTTCAATGATACACACCAAGGTACTACT	120	TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster	1	FILE REFERENCE: File Reference: 7326-094	
QY	2759	GTGTTTGACACGCTACCGAATCTGTTGTTCAACAAATGGGACCGATCTACTAATTCAT	2818	CURRENT APPLICATION NUMBER: US/09/270,767	1	CURRENT FILING DATE: 1999-03-17	
DB	121	GTGTTTGACACGCTACCGAATCTGTTGTTCAACAAATGGGACCGATCTACTAATTCAT	178	NUMBER OF SEQ ID NOS: 62517	1	SOFTWARE: PatentIn Ver. 2.0	
QY	2819	CTGTCATCTGTTGTTGACGCGGTTGTTCAACAAATGGGACCGATCTACTAATTCAT	2818	SEQ ID NO 596	1	LENGTH: 767	
DB	179	CTGTCATCTGTTGTTGACGCGGTTGTTCAACAAATGGGACCGATCTACTAATTCAT	237	TYPE: DNA	1	ORGANISM: Drosophila melanogaster	
QY	2879	GCAACCAAGATTGGAAGAGCTACAGGTGTCAGAAATATTTCCACGAGATTGCA	2938		1		
DB	238	GCAACCAAGATTGGAAGAGCTACAGGTGTCAGAAATATTTCCACGAGATTGCA	297		1		
QY	2939	GCTTAAAGAACTATTAGATCATATGATTTCTACTCAAAAGCACTTCAAGAGCG	2998		1		
DB	298	GCTTAAAGAACTATTAGATCATATGATTTCTACTCAAAAGCACTTCAAGAGCG	357		1		
QY	2999	CATGCCATTCTGCGCTTCTTACGATCTGCGCAATACATCTGTTAAAGCAGCACTT	3058		1		

```

Db      |||
358  CATGCCATTCTGCGCTTCTTACGATCTGGCCAAATACATCTGTAAAGACGCACTT 417
Qy      |||
3059  TCCACGACACGAGCCAGGCGGCGAGGAGGAGCGGCTTAAAGCTGGCCAGTGGCGG 3118
Db      |||
418  TCCACGACACGAGCCAGGCGGCGAGGAGGAGCGGCTTAAAGCTGGCCAGTGGCGG 477
Qy      |||
3119  GGATCAGAGTGAATGAGAGTGGCTCCAGGCGGAGAGTGGCTTAAAGCTGGCGG 3178
Db      |||
478  GGATCAGAGTGAATGAGAGTGGCTCCAGGCGGAGAGTGGCTTAAAGCTGGCGG 537
Qy      |||
3179  CCTGAATACCCCGCAATGACCTTTTACTGCGGCACTGGCCAAAGTATATCAGAGCTC 3238
Db      |||
538  CCTGAATACCCCGCAATGACCTTTTACTGCGGCACTGGCCAAAGTATATCAGAGCTC 597
Qy      |||
3239  TGCCCGATCTGCAATCCATTCGACATGAGGCGGCAAGCCAAAGTGAATTCGGGTCG 3298
Db      |||
598  TGCCCGATCTGCAATCCATTCGACATGAGGCGGCAAGCCAAAGTGAATTCGGGTCG 657
Qy      |||
3299  CAATTGCGCTAGTACTCTCCATTGAGCAGTAGCAGTCCCAAGCTTTGCAAGTCCCTC 3358
Db      |||
658  CAATTGCGCTAGTACTCTCCATTGAGCAGTAGCAGTCCCAAGCTTTGCAAGTCCCTC 717
Qy      |||
3359  TGTGGGTCCATTGCTGCTTCCAGAGCCACCAACCACTTAGCATCA 3408
Db      |||
718  TGTGGGTCCATTGCTGCTTCCAGAGCCACCAACCACTTAGCATCA 767

```

RESULT 2
US-09-270-767-15878
Sequence 15878, Application US/09270767
Patent No. 6703491

GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 15878
LENGTH: 767
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-15878

Query Match 14.2%; Score 732.8; DB 3; Length 767;
Best Local Similarity 98.7%; Pred. No. 8e-208;

Matches 760; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

```

Qy      |||
2639  CCTGATCAGCTTCAACAGATATCTGCGGATGAGAAAGAGCTCACTCAGACGAATTA 2698
Db      |||
1  CCTGATCAGCTTCAACAGATATCTGCGGATGAGAAAGAGCTCACTCAGACGAATTA 60
Qy      |||
2699  ATGCGTTCTGACATGTTCCGCAAGTGTTCATGATACACACCAAGGTGTACTACT 2758
Db      |||
61  ATGCGTTCTGACATGTTCCGCAAGTGTTCATGATACACACCAAGGTGTACTACT 120
Qy      |||
2759  GTTTTGGACAGGTCACCGAACTGATTTCTGTTATGCGAAGAACTTACAGAAATGG 2818
Db      |||
121  GTTTTGGACAGGTCACCGAACTGATTTCTGTTATGCGAAGAACTTACAGAAATGG 178
Qy      |||
2819  CTCTTCATCTTGTGTTGACGCGGTGTTCAACAATTTGGGACCGGATCTACTCAATTCGAT 2878
Db      |||
179  CTCTTCATCTTGTGTTGACGCGGTGTTCAACAATTTGGGACCGGATCTACTCAATTCGAT 237
Qy      |||
2879  GCACAGCAAGATTTGGAAGACGCTACAGGTGTCCACGAATATTTCCACGCAAGTTGA 2938
Db      |||
238  GCACAGCAAGATTTGGAAGACGCTACAGGTGTCCACGAATATTTCCACGCAAGTTGA 297
Qy      |||
2939  GCTTAAAGAACTATTAGAAATATATGGAATTTCTACTCAAGGCCAACTACCAAGCGCG 2998
Db      |||
298  GCTTAAAGAACTATTAGAAATATATGGAATTTCTACTCAAGGCCAACTACCAAGCGCG 357

```

```

Qy      |||
2999  CATGCCATTCTGCGCTTCTTACGATCTGGCCAAATACATCTGTAAAGACGCACTT 3058
Db      |||
358  CATGCCATTCTGCGCTTCTTACGATCTGGCCAAATACATCTGTAAAGACGCACTT 417
Qy      |||
3059  TCCACGACACGAGCCAGGCGGCGGAGGAGGAGCGGCTTAAAGCTGGCCAGTGGCGG 3118
Db      |||
418  TCCACGACACGAGCCAGGCGGCGGAGGAGGAGCGGCTTAAAGCTGGCCAGTGGCGG 477
Qy      |||
3119  GGATCAGAGTGAATGAGAGTGGCTCCAGGCGGAGAGTGGCTTAAAGCTGGCGG 3178
Db      |||
478  GGATCAGAGTGAATGAGAGTGGCTCCAGGCGGAGAGTGGCTTAAAGCTGGCGG 537
Qy      |||
3179  CCTGAATACCCCGCAATGACCTTTTACTGCGGCACTGGCCAAAGTATATCAGAGCTC 3238
Db      |||
538  CCTGAATACCCCGCAATGACCTTTTACTGCGGCACTGGCCAAAGTATATCAGAGCTC 597
Qy      |||
3239  TGCCCGATCTGCAATCCATTCGACATGAGGCGGCAAGCCAAAGTGAATTCGGGTCG 3298
Db      |||
598  TGCCCGATCTGCAATCCATTCGACATGAGGCGGCAAGCCAAAGTGAATTCGGGTCG 657
Qy      |||
3299  CAATTGCGCTAGTACTCTCCATTGAGCAGTAGCAGTCCCAAGCTTTGCAAGTCCCTC 3358
Db      |||
658  CAATTGCGCTAGTACTCTCCATTGAGCAGTAGCAGTCCCAAGCTTTGCAAGTCCCTC 717
Qy      |||
3359  TGTGGGTCCATTGCTGCTTCCAGAGCCACCAACCACTTAGCATCA 3408
Db      |||
718  TGTGGGTCCATTGCTGCTTCCAGAGCCACCAACCACTTAGCATCA 767

```

RESULT 3
US-09-270-767-6018
Sequence 6018, Application US/09270767
Patent No. 6703491

GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 6018
LENGTH: 1420
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-6018

Query Match 12.8%; Score 661; DB 3; Length 1420;
Best Local Similarity 99.0%; Pred. No. 3.4e-186;

Matches 676; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

```

Qy      |||
4465  TGTATATGAAAGAGGAGGAACTGCAATCAGTGGGAGGAAAGCTTCAAAAGACTCCT 4524
Db      |||
1  TGTATATGAAAGAGGAGGAACTGCAATCAGTGGGAGGAAAGCTTCAAAAGACTCCT 60
Qy      |||
4525  CCGCGGATCGTCTGATGATTTGGGAGAGCCCTTATAGATTCCTGC--TCGTGCCACC 4582
Db      |||
61  CCGCGGATCGTCTGATGATTTGGGAGAGCCCTTATAGATTCCTGCATACGACACACC 120
Qy      |||
4583  ACAAAACAAGACACAGACGCGGTGCTTCCCTGCGCTGAGAGAGAGTGAAGAGCGGCG 4642
Db      |||
121  ACAAAACAAGACACAGACGCGGTGCTTCCCTGCGCTGAGAGAGAGTGAAGAGCGGCG 180
Qy      |||
4643  ACATTAATATATATATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4702
Db      |||
181  ACATTAATATATATATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240
Qy      |||
4703  CTTTATCTGTATGATCTGATGATGATATTTATTCAGACAGATCAATTCGCGGCGGAA 4762
Db      |||
241  CTTTATCTGTATGATCTGATGATGATATTTATTCAGACAGATCAATTCGCGGCGGAA 300
Qy      |||
4763  AGTCAAGGCGCATGCGCATCTAATGTGAGCAAGAAAACCTATTTATACATAACGGGAA 4822

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 23, 2005, 14:59:23 ; Search time 54 Seconds

(Without alignments)
2658.434 Million cell updates/sec

Title: US-10-030-850-2

Perfect score: 7611

Sequence: 1 MAYRRPSDDGFIQWPKAD.....RNCISGGSSSTKNSNASSS 1492

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 263416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1375	18.1	1289	2 T00387	KIAA0622 protein -
2	1348	17.7	1324	2 T00386	hypothetical prote
3	611	8.0	1063	2 E88546	protein R107.6 (im
4	593.5	7.8	1080	2 S30876	hypothetical prote
5	389	5.1	826	2 C88550	protein ZC84.3 (im
6	328	4.3	1972	2 S68176	TOG protein - huma
7	277.5	3.6	730	2 S28294	hypothetical prote
8	262.5	3.4	643	2 S28293	hypothetical prote
9	259.5	3.4	1462	2 T11658	probable mitotic s
10	198.5	2.6	5327	2 T13564	microtubule-associ
11	198	2.6	1148	2 F86403	probable transposo
12	195	2.6	199	2 C84586	hypothetical prote
13	195	2.6	2271	2 F90073	hypothetical prote
14	193	2.5	2232	2 T34434	hypothetical prote
15	192	2.5	4377	2 A55575	ankyrin 3, long sp
16	190	2.5	860	2 T39502	hypothetical prote
17	184	2.4	1032	2 T34433	hypothetical prote
18	182.5	2.4	1955	2 T30934	myosin-like protei
19	180.5	2.4	1630	2 A53577	ascies sialoglyco
20	180	2.4	809	2 T41645	probable spindle p
21	179.5	2.4	1513	2 T50073	mitotic spindle pr
22	179	2.4	1727	2 T50073	myosin-like coiled
23	177.5	2.3	3507	2 T34513	hypothetical prote
24	177	2.3	1108	2 T17455	translation initia
25	177	2.3	2346	2 T00057	hypothetical prote
26	175.5	2.3	2346	2 T13829	typ homolog - frui
27	174.5	2.3	2021	2 A84771	zys-9 protein - Ca
28	173	2.3	1415	2 T21244	serine/threonine-s
29	172.5	2.3	1114	2 T14351	

30	172	2.3	725	2 A41258	a-agglutinin core
31	172	2.3	1104	2 S5310	probable membrane
32	172	2.3	3225	2 I52300	giantin - human
33	172	2.3	3259	1 A56539	giantin - human
34	171	2.2	927	2 T38127	phosphoprotein - f
35	171	2.2	1326	2 AC3372	kinasin-like prote
36	170.5	2.2	2481	2 D90011	hypothetical prote
37	170	2.2	1792	2 T20363	hypothetical prote
38	169.5	2.2	6713	2 B85921	hypothetical prote
39	168	2.2	952	2 T18837	hypothetical prote
40	168	2.2	1912	2 T29088	vitellogenin I pre
41	167.5	2.2	1520	2 T44231	hypothetical prote
42	167	2.2	1237	2 T45070	protein kinase hom
43	167	2.2	2012	2 T43214	ovtl protein - nem
44	164	2.2	2819	2 A90551	conserved hypotet
45	163.5	2.1	1072	2 A86827	hypothetical prote

ALIGNMENTS

RESULT 1

KIAA0622 protein - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004

C/Accession: T00387

Rishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, DNA Res. 5, 169-176, 1998

A/Title: Prediction of the coding sequences of unidentified human genes. X. The comple

A/Reference number: Z14142; M01D:98403880; PMID:9734811

A/Accession: T00387

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-1289 <158>

A/Cross-references: UNIPROT:O75118; UNIPARC:UPI000007259A; EMBL:AB014522; NID:93327057

A/Genes:

A/Note: KIAA0622

Query Match 18.1%; Score 1375; DB 2; Length 1289;

Best Local Similarity 30.1%; Pred. No. 5,76-69;

Matches 416; Conservative 235; Mismatches 495; Indels 234; Gaps 46;

QY	236	PSALKNTNGVGLDEADNIGLRERPTMKPLHSAVSSLRPKPNVDVTDGAGATM	295
DB	15	PSRRN-----VGMGTRRLG-----	50
QY	296	ESFESSFEVVPQNTFPHADMDIYKQVLTISDKNADMEKVDALKTRALLISYHQ	355
DB	51	EDFIKAFDDVPVQVYSSRDLRESINKIREILSDDKHDEORVNAIKTRISLLAGAAEY	110
QY	356	PQVAVQLELSFVDILKEELRSQVIREACTITTAVMSKTLRNKIDAFQMSILEHLNL	415
DB	111	DNFQO-HLELLDGA-F-KLSAKDLRSQVVEACTTGLHLSVLGNKNDHGAEMIPITFNL	168
QY	416	IQNSAKVIASASTALIKYIKYTHAPKLIKITYDTLNOSKSDIRSTLCMLVLFEEWG	475
DB	169	IPNSAKIMATSGVAVRLIRHTHPRLLPVITSNT-SKSAVARRCFEFDLLQEWQ	227
QY	476	TYALEFNATVLDTLKKSIGDADCDARRSRAYAFRRHFPFLADQITGLDIAQRL	535
DB	228	THSLERHISVLAETIKKIGHDADSEARIFARKCYWGFHSHFSEAEHLVHTLESSYOKAL	287
QY	536	ENREBEGGGGGTGTGTGTAPEPTRRYSRIGTPGTQKTPPSMRISAVDTAAQAQYR	595
DB	288	QSHLK-----NSDSIVSLPQSDRSS-----SSQESINRPLSAKRSFTGTSASIVSTK	338
QY	596	AQYTLVSRQKPLGPNNSQAASGTAAGSLPRPLNSNGT-----ATPGSV---	647
DB	339	SVSTTSLQ-----SRSDIDVNAASA-----SKSVSSSSGTTTPSSAAALPPGYSASL	368
QY	648	-----TPRPRGAG-----VQSO-----PSRSSTSPETKLRDQY	677

Db 389 GRIITRRQSGSNTVASTPDNKRGRARVAVSQSQRSGSANPAGAGSRSSSPGKLIGSY 448
 Qy 678 GGI -GNVYRG -ATGATPKKASGIPRSTASRETSPTRSQ---GGLMKRMYSTGAG---S 729
 Db 449 GGITGGSSSGPPVTPSESKRSKIPKQSGSRETSPTRIGLANSSRIPRSMQCGSRDTS 508
 Qy 730 RRTPERNNVPRSPAPARLLAQSREAHETLGVGDQDQ-PPVSGDYNRSGG----- 778
 Db 509 RSSNRDTPARGPPL-----DRGLGQPGRIIPGSVANMRVLSTLDLEAVADA 558
 Qy 779 ---MRMRKLMGRDESDDIIDEASVCSERSPDSYTTGNKSNVSLSGSHRLDMSTOR 834
 Db 559 LKKPVRRRYEPYGMWSDDANSSVCSERSYGS---RNGGIPHY-----L 602
 Qy 835 APFDIETLIIQCASTHMERKXGGLSLNQYADGKELQOQCKYLDIMFKKPMQDTHK 894
 Db 603 ROTEDVAEVLNHCASSNMBERKGLGLQNLKSGRTLSRVELKICEIFTRFADPHK 662
 Qy 895 -VYSLFLDTVTTELILVHANETSNGSSCLTRLFNKLTGDLNLSMHSKIMKTLOVVEYF 953
 Db 663 RYFSMFLTELVDPIIHKDL-QDWLFVLLTQLKMGADLLGSVQAKYQKALDVTBDSF 721
 Qy 954 PPOLOKELFRIISDSTOTPTTKTIAIRFLTDLANTYCKSSDPFSD--QSOACERTVL 1011
 Db 722 PPDQOFNIMRPFVDTOTPNLKVKVAILIKYTESLA---RQWD-PTDFVNSSETRLAVS 776
 Qy 1012 KLAQLAADQKSMELRQARSCLVALYNLTPQWTLTLADLPKYQDSARSCHSHMRQS 1071
 Db 777 RIITWTEKSSDVRCAAQIVLISPELTPERTMLGALPKTPQDGAIKLHLNKLNS 836
 Qy 1072 QSCNSGANSPPS-----SPLSSSPKPLQSPS-----VGPPA-----SLQSHHOL 1112
 Db 837 ---NNSVGSPTMTIGTRPSRHTSSRTSPILSPNCSHGGLSPRLMGWGSADGLAKPRPF 893
 Qy 1113 STGSSTPRSRQSSVEOELLFSSSLDIQ-HNIQKTSSEIHCHGCGQYQTLAPNGRHHQ 1171
 Db 894 SQPNSIPTAPSHKALRSYSPSLDYDTENTL--NSEIYISLQVTE-AIEKFSFSGQD 950
 Qy 1172 YHDOGGOD---SCASLSNSKTQSSANTTQSNTP--ESATMBRLDNLERETTONAKS--- 1223
 Db 951 LNEPIKRDGKKECDIYSRQGAASPAKTEGSGSEVGGRTALDNKTSILANTOPPRAPFPP 1010
 Qy 1224 -----PTDDAKVITVSINNAENGEL-----ILASNMISEV 1254
 Db 1011 RARDVNPYPSDA-----INTYDKALKEAVFDDMEQLRDVPIDHSDLVADLKELSN 1064
 Qy 1255 VVALTLTKDQVELLQ-TSLTNLGCICIGANCELPKNGFRSIRMRLNLIILAEHTDVI 1313
 Db 1065 HNERVEERKALLETIKITREDSLGW-----EEHFKTILLLLTLLDGDKHSIRA 1115
 Qy 1314 AGLHYLSKIMRSNKRHMNHFLLELTKIQCQHS-KEALRDIDSMIPRIAPSLPLD 1372
 Db 1116 LALRYARETLRKNOPAF--EKNYAELTITMKTLEHAKSHKEVRAAEAASTLASSIHPEQ 1173
 Qy 1373 SINIVNVIATGEBFTNLCAIKILLTEVTHGSEITDAHLDIVFPNLAASADDTOSWAK 1432
 Db 1174 CTKVLCPIITQADYPIINLAIKMQTVVERIAKESLLQLLVDIIPELGCVNTBESYRK 1233
 Qy 1433 AAVFCVLYKLYLGEKVKPKSVLNPSTKRLNNTYIEKORNCISGGSGSTKNSAAGS 1492
 Db 1234 ASVFLVALIYSVIGED-LKPHLAQLTGSKKLLNLYIKR-----AQTNNSSSSS 1283
 RESULT 2
 T00386
 hypothetical protein KIAA0627 - human (fragment)
 C1:Species: Homo sapiens (man)
 C1:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
 C1:Accession: T00386; T46487
 R1:Shikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Koetani, H.; Nomura, N.
 DNA Res. 5, 169-176, 1998
 A1:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
 A1:Reference number: Z14142; M0ID:98403880; PMID:9734811

A1:Accession: T00386
 A1:Status: preliminary; translated from GB/EMBL/DBJ
 A1:Molecule type: mRNA
 A1:Residues: 1-1324 <18>
 A1:Cross-references: UNIPROT:075122; UNIPARC:UPI000006F360; EMBL:AB014527; NID:G3327067
 A1:Experimental source: Brain; clone H01026
 A1:Due to: Hecht, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A1:Reference number: 223035
 A1:Accession: T46487
 A1:Status: preliminary
 A1:Molecule type: mRNA
 A1:Residues: 969-1324 <AAA>
 A1:Cross-references: UNIPARC:UPI000016ACCA; EMBL:AL137636
 A1:Experimental source: adult testis; clone DKFZp434F075
 C1:Genetics:
 A1:Note: KIAA0627; DKFZp434F075.1
 Query Match 17.7%; Score 1348; DB 2; Length 1324;
 Best Local Similarity 29.6%; Pred. No. 2e-67;
 Matches 404; Conservative 233; Mismatches 523; Indels 204; Gaps 39;
 Qy 237 SALK-----NTNGNGVGLDEADNIGURERTMIKREPLHSAYSSSLRPRPNVDVTDGA 292
 Db 55 SAFKVPARKTSGN-----PANSARKP-----GSAGPVGASKEGAGGA 94
 Qy 293 VTMESEFSEFVFPOLNIFHAKMDMDIYKOVLIISDGNADMEKRVDAIKTIRALLI--- 349
 Db 95 VDEDFITAFIDVDSIQYSSRELEETLNKIREISLSDKHMDQANLKKIRSLILVAGA 154
 Qy 350 LSYHTQPOFVAVOLKELSLSFVDILKEELRSQVIREACITTAIYNSKTLNKLDAFCWSIL 409
 Db 155 AQYDFEFOHRLDLQALKLS-----AKDLRSQVAREACTIAHLSTVLGNFEDHGAEAIV 209
 Qy 410 EHLNLIONSANVIVASASTIALKTIKTYHAPKLIKITYTDLNOSKSDISTLCEMLVL 469
 Db 210 PTLFVLVNSAKVATSGCAIRFIIRTHVPRILPLITSNCT-SKSVVRRRSFEPIDL 268
 Qy 470 LFEENQATLRNNAITVLEDTLKKSIGDADCDARHRSRYAVAFRRHPPELADQIYGLDI 529
 Db 269 LIGEQHTSLERHAAVLVETIKKGIHDDADARVARKTYGRLNHFGEAETLYNSLEP 328
 Qy 530 AAQRALEEREREGGG-----GGTGTGTAPETRTVSR 564
 Db 329 SYQSKLQYLYLKSSGVSASLPQSDRSSSSSQESLNRPFSSKMTANPSTAVRVSAGSSKA 388
 Qy 565 GRTBCTLOKPTPMSRISAVDTAAQRAKVAQVTLVRSQKPKPGNNNSQAS----- 617
 Db 389 SLPESLQ-----RSRSDIIVNNAAGAKAH-HAAGQVRRGRUGAGALNAGTASLEDT 441
 Qy 618 ---MTGAAGSLPFRPLNSNSGTPATTPGSAVTPRPRGAG---VSQSGPGRSTSP-- 669
 Db 442 SDKLDGTASEQGRVAKLSA-----PLAGMNAADSGRGRRTKTVNSQSGSRGSGR 496
 Qy 670 ---STKLDYQGGIGNYRGATGAI PKKASGIPSTASRETSPT---SGGGLMKRMY 723
 Db 457 VLTTLALSTVSGVQVVLNLSASA--QKRSKIPSGGCSREASPSRLSVARSRRIPRSV 554
 Qy 724 STGA---GSRTPERNNVPYRSP-----ARLAAQ 750
 Db 555 SQGCSREASRSDTSRVSFQPLASRHSRSTGALYAPVYVYAGSGGYIGSQSRSLSS 614
 Qy 751 SREAEHTLGVDDGQPDVYSGDYNRSGGMGRKLMGRDESDDIIDEASVCSERSFPDS 810
 Db 615 SVSARVAVNTGSD--VEEAVADALKKPRARR--YESYGHSDDDANSDASSACSERSYS- 670
 Qy 811 YTRGNKSNVSLSGSHTRLDMSTQAPPDIDITIIQFCASTHMERKXGGLISLTQYLAQK 870
 Db 671 -----RNGSI-----PTYMRKOTEDVAEVLNRCASSWMSKKEBLLQNLKOR 715
 Qy 871 ELTQOQKLCVDMFRKMFNDTHK--VYSLFDYDTTELILVHANETSNGSSCLTRLFNK 929
 Db 716 TLSRVELKRLCEIFTRFADPHGKRVFSWFLTELVDIYVHKDL-QDWLFVLLTQLLKK 774

Db 361 ATCGAAGTACTACTTACGAGGAGCCACCGCGCTTGAGCTGAGACAGCCAGCGGCGCC 420
 QY 2573 ATTGACGACATGAGACGATATTTCAGTTCTGGCGATGAGCAT 2618
 Db 421 ATTGACGACATGAGACGATATTTCAGTTCTGGCGATGAGCAT 466

RESULT 6

US-10-104-047-327
 Sequence 327, Application US/10104047
 Patent No. 6943241
 GENERAL INFORMATION:
 APPLICANT: HELIX RESEARCH INSTITUTE
 TITLE OF INVENTION: No. 6943241el full length cDNA
 FILE REFERENCE: H1-A0105
 CURRENT APPLICATION NUMBER: US/10/104,047
 PRIOR FILING DATE: 2002-03-25
 NUMBER OF SEQ ID NOS: 4096
 SOFTWARE: Patent Ver. 2.1
 SEQ ID NO 327
 LENGTH: 3153
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-104-047-327

Query Match 2.9%; Score 149.8; DB 3; Length 3153;
 Best Local Similarity 47.1%; Pred. No. 2.9e-33;
 Matches 685; Conservative 0; Mismatches 732; Indels 36; Gaps 6;

QY 232 CATGCTGACGCGGACCCATTAAAGATTGCACAAAAAGTCCCTGGAGCGTTCTCGAGC 291
 Db 521 CTGGGCTGAACCTTAGCAATTACAAAGTGTCTGCTGGAGAGACATCTGTCGCC 580
 QY 292 TAATTAAGCATTTGGGCGAGCATTTTAATGCATACACGCTACCGTTCTGCCACATGTA 351
 Db 581 TGGTACCCTGGAGATCGGATCGTTTCAAGCGCAGATGCGACATGCTGCTCAAGTCTTA 640
 QY 352 TCGATCGCTGGAGACAGACAGGACACAGTCCGCGAGAGCGCACTTCTGCTGGCGG 411
 Db 641 TAGACAGACTAGAGATGCTAAAGACTGTGAGGAGAGAGACCAAACTGCTGCTTAA 700
 QY 412 ACCTGAGGACAGAGTCTTCCGCCCGCCAGCGCTGATCGACAAGCTGGCCACTAGCT 471
 Db 701 AANTCATGATCAAGCTGCTAATCCCACTAGCTATGAGACAGATGCT-----TGAG 754
 QY 472 GCTTCAAGACAAAGACGCGAAGTGGCGAGAGTCTTTCAGACGATTGTGAACGCTG 531
 Db 755 GCTTCAAAACAAAGATTTTCGTACTGAGAGAGCATCTGCTGCTTATGCAACAC 814
 QY 532 TTGATGATGCGGACCCAGCACTTATGTTGCGCTTATTAACCAACAGTTTGTGCAC 591
 Db 815 TCAATGCTTGGAGACACAGACTTAACTTAAGCAAGATTGTCACATATATGCACT 874
 QY 592 TTCTCGAGATCCCACTTAATGTGAGGAGGCGCATCCAAAGCTAGTGAATCT 651
 Db 875 TACTTGAGATCCAAACAGCAGGTTGAGATGAGCAATTAACAGTTAGTGAATTT 934
 QY 652 ACAAGCATTAAGGAGATGATGAGGCGCAGCTCCGTCGATGAGCATGTTCTGCT 711
 Db 935 ACAACATTAAGGAGAGCTGTGAGGCGCAGATCTCAGTAA-----AAGATTGCGACAGT 991
 QY 712 CGAATTGCTATGTTGAGCAAAAGTTGACCGAGTCAACAGAGAGGTCTACTGCTAC 771
 Db 992 CCGGTTGAATGATTTTACAAATTTGATGAATCCAAATCTGGAACATGATAC 1051
 QY 772 CTGAGCCCTTAAACAGATGAGATGAGGCGCTTGACAGAGCC-GACATATT 830
 Db 1052 AATCGCAATGATTAATAATTTGACGATGAAGATTTCTGTGATGTAACAGACCTTCT 1111
 QY 831 GGGTTGAGGAGGACCCAGCATGATTAAGCGGCACTAAGCTGCGCGTTTGTGCA 890

Db 1112 CTGCTAGTCTACATCAATCCAGGCTCCACCAGTTCTCGAGAAAGATTGGAATGGAA 1171
 QY 891 -----TCACTGCGCCCAAAACCAATGTGAACGATGTACCGGTATG 933
 Db 1172 CCACCCCGCGCTGTGTTCACTCCACCTTGATCCAAAGTCTTCAGCTGCCAAAGAGAG 1231
 QY 934 CCGGCGCGTAAACAGTGAATCTTTGATCTAGCTTTAGAGTGTCCGCAATGGAACA 993
 Db 1232 CTGCTGCTGTGATGAAGATTTTATTAAGCATTTGATGATTAAGTCTGTAACAGA 1291
 QY 994 TCTTCCAGCTTAAGACATGAGCATATCTCAACAGCAAGTACTAGTATCATCACTGATA 1053
 Db 1292 TTTATTCAGCGGACCTTGAGGATCTATTAACAAATTAAGGAAATATATCTGATG 1351
 QY 1054 AAAAGCAGACTGGAGAAACGTGTGATGCTCTCAAGAAATCGAGGCTTGTCTATTC 1113
 Db 1352 ACAGCATATGTTGAGAGCAGAGATTAATGCTTAAAGATTAAGATTTTACCTTT-- 1409
 QY 1114 TCAGCTATCACACTGACCGCAGTTGTGCTGTACAGCTAAAGAAATGTGTTAAGCT 1173
 Db 1410 ----GGCTGTGCTGTGATGATTAATCTTTTCAACATTTGCTTTTGGATGAG 1465
 QY 1174 TCGTGAATCTCTCAAGAGAGAACTAGATCAAGGTGATCCGCGAGCGTGATCAACA 1233
 Db 1466 CTTTAACTCTCTGCTAAGACCTGCGGCTTCAGGTAAGTGGAGGCTGTATCAAGT 1525
 QY 1234 TCGCTAATCTGTAGACGCTGAGAAATTAAGTATGCTTCTGCTGAGCATTTTG 1293
 Db 1526 TGGGAGTCTGTATCATGTTCTGAGGATTAAGTTGACATGAGCTGAAGCATTAATGC 1585
 QY 1294 AGCAGCTGATTAATTAATACAGAACGCGGAAAGCTCATTTGATCCGCTTCAATAG 1353
 Db 1586 CAACATCTTTAATTAATTCAAACAGTCCAAATTAAGCCACATCTGCTGTAG 1645
 QY 1354 CTGTAAGTATATCATTAATTAATACATGACCAAGCTGCTCAAGATCTACAGACA 1413
 Db 1646 CTGTTAGTTAATTAATTTGAGACACACATCTTATGTTAATACGTCTAACAAGCA 1705
 QY 1414 CTGTAATCAATCAAGTTCGAAGACATTAAGTCCACATCTGTGAGTGAATGCTGCG 1473
 Db 1706 ---ACTGACTCTTAAGTCTGTGCAATTAAGAGCGCTGTTTGAATTTTGAATTTGC 1762
 QY 1474 TCTTGAAGATGCGACAGAGCGTTTGAAGAAATGCCAGCTATCAAGGACACT 1533
 Db 1763 TTTTAAAGATGAGAGACATCTACATGAACGACATATACATTAAGTGAACAA 1822
 QY 1534 TAAATAATCATTTGCGATGACAGCTGCGATGACGCGCATTTCAAGATGCGCTATT 1593
 Db 1823 TAAAGAGGAATACATATGCTGATTCGGAAGCAAGAAATGAACAGAAATGTTACT 1882
 QY 1594 GGGCTTTGAGGCTGCTTTCCAGAGCTGGCGGATCAATATATGGAACATTAGACATG 1653
 Db 1883 GGGGTTTCCAGTCACTTCCAGAGAGAGAGACACATCTTGAACACACCTTGGAGTCT 1942
 QY 1654 CTGCCAGGCGC 1666
 Db 1943 CCTACCAAGAAC 1955

RESULT 7

US-09-270-767-30401
 Sequence 30401, Application US/09270767
 Patent No. 6703491
 GENERAL INFORMATION:
 APPLICANT: Homburger et al.
 TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 FILE REFERENCE: File Reference: 7326-094
 CURRENT APPLICATION NUMBER: US/09/270,767
 CURRENT FILING DATE: 1999-03-17
 NUMBER OF SEQ ID NOS: 62517
 SOFTWARE: Patent Ver. 2.0
 SEQ ID NO 30401

Db 301 AGTAAAGGCGCATGCCCATTCTAATGTAGCAGAAAACTATTATACATAAAGGGAAC 360
Qy 4823 TAAAGCAGTAATAATCCGCAAGTGTACTAAATTTGAGCCAACTCCGCAATCTCACTTC 4882
Db 361 TAAAGCAGTAATAATCCGCAAGTGTACTAAATTTGAGCCAACTCCGCAATCTCACTTC 420
Qy 4883 TCACATCAGTTGACGTCTTTTACAGCTCTTAATTAACGCTGATTCGTTTAAT 4942
Db 421 TCACATCAGTTGACGTCTTTTACAGCTCTTAATTAACGCTGATTCGTTTAAT 480
Qy 4943 TGTAAAGCTATATACCGCTCTTTATGAGAACTAGAGAGTGTCTTTCTACCTTTGTTTC 5002
Db 481 TGTAAAGCTATATACCGCTCTTTATGAGAACTAGAGAGTGTCTTTCTACCTTTGTTTC 540
Qy 5003 AGGCGACTTGTATCCCTTATTTCCGCTCACTAAATTTCTAATTAATTAAT 5062
Db 541 AGGCGACTTGTATCCCTTATTTCCGCTCACTAAATTTCTAATTAATTAAT 600
Qy 5063 AGCATTTCTGCTTTACACTGAACCTATTTTGTAAATTAATTAATTAATTAAT 5122
Db 601 AGCATTTCTGCTTTACACTGAACCTATTTTGTAAATTAATTAATTAATTAAT 660
Qy 5123 AAAGATTAATAATCAATTAATAA 5145
Db 661 AAAGATTAATAATCAATTAATAA 683

RESULT 4
US-09-270-767-21300
Sequence 21300, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 21300
LENGTH: 1420
TYPE: DNA
ORGANISM: *Drosophila melanogaster*
US-09-270-767-21300

Query Match 12.8%; Score 661; DB 3; Length 1420;

Best Local Similarity 99.0%; Pred. No. 3,4e-186;

Matches 676; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

Qy 4465 TGTACATCAGAAAGCAGCGGAATCGATCAGTGGGAGAGAGCTTCAAGAACTCCT 4524
Db 1 TGTACATCAGAAAGCAGCGGAATCGATCAGTGGGAGAGAGCTTCAAGAACTCCT 60
Qy 4525 CCGCGGCATCGTGTGATGATGCGGAGCCCTTAATAGATTCCTGC--TCGTGCACC 4582
Db 61 CCGCGGCATCGTGTGATGATGCGGAGCCCTTAATAGATTCCTGCATCAGACACACC 120
Qy 4583 ACAAACAAGACAGACCGGGTTCCTTCCCTCGGCTGAGAAAGAGTGAAGAGCGGCG 4642
Db 121 ACAAACAAGACAGACCGGGTTCCTTCCCTCGGCTGAGAAAGAGTGAAGAGCGGCG 180
Qy 4643 ACATTAATAATATATATTTATTAATTAATTAATTAATTAATTAATTAATTAAT 4702
Db 181 ACATTAATAATATATATTTATTAATTAATTAATTAATTAATTAATTAATTAAT 240
Qy 4703 CGTTACTGTGTACTGCTCGATGATATATTCAGAGATGCAATGCGGCCGAGAAAG 4762
Db 241 CGTTACTGTGTGTACTGCTCGATGATATATTCAGAGATGCAATGCGGCCGAGAAAG 300
Qy 4763 AGTCAAGGCGCATCGGCCATCTTAATGTGAGCAAGAAATCTATTTATTAATTAAT 4822
Db 301 AGTCAAGGCGCATCGGCCATCTTAATGTGAGCAAGAAATCTATTTATTAATTAAT 360

Qy 4823 TAAAGCAGTAATAATCCGCAAGTGTACTAAATTTGAGCCAACTCCGCAATCTCACTTC 4882
Db 361 TAAAGCAGTAATAATCCGCAAGTGTACTAAATTTGAGCCAACTCCGCAATCTCACTTC 420
Qy 4883 TCACATCAGTTGACGTCTTTTACAGCTCTTAATTAACGCTGATTCGTTTAAT 4942
Db 421 TCACATCAGTTGACGTCTTTTACAGCTCTTAATTAACGCTGATTCGTTTAAT 480
Qy 4943 TGTAAAGCTATATACCGCTCTTTATGAGAACTAGAGAGTGTCTTTCTACCTTTGTTTC 5002
Db 481 TGTAAAGCTATATACCGCTCTTTATGAGAACTAGAGAGTGTCTTTCTACCTTTGTTTC 540
Qy 5003 AGGCGACTTGTATCCCTTATTTCCGCTCACTAAATTTCTAATTAATTAAT 5062
Db 541 AGGCGACTTGTATCCCTTATTTCCGCTCACTAAATTTCTAATTAATTAAT 600
Qy 5063 AGCATTTCTGCTTTACACTGAACCTATTTTGTAAATTAATTAATTAATTAAT 5122
Db 601 AGCATTTCTGCTTTACACTGAACCTATTTTGTAAATTAATTAATTAATTAAT 660
Qy 5123 AAAGATTAATAATCAATTAATAA 5145
Db 661 AAAGATTAATAATCAATTAATAA 683

RESULT 5
US-09-270-767-14266
Sequence 14266, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 14266
LENGTH: 466
TYPE: DNA
ORGANISM: *Drosophila melanogaster*
US-09-270-767-14266

Query Match 8.8%; Score 452.4; DB 3; Length 466;

Best Local Similarity 99.6%; Pred. No. 3,2e-124;

Matches 464; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 2154 GGAATACCCCGAAGCAGCAGCTCTAGGAAAGCAAGTCCAATAGTCAAGTGTGCT 2213
Db 1 GGAATACCCCGAAGCAGCAGCTCTAGGAAAGCAAGTCCAATAGTCAAGTGTGCT 60
Qy 2214 TTGATGAACGAGATGTACTCTACAGTGGGAGTCTCGACGTACGCCGAGAGAAC 2273
Db 61 TTGATGAACGAGATGTACTCTACAGTGGGAGTCTCGACGTACGCCGAGAGAAC 120
Qy 2274 AACCCAGTAAGCAATCGGCGCGGACAGCA--CTGCTGGCGCAATCCGTAAGCAGAA 2332
Db 121 AACCCAGTAAGCAATCGGCGCGGACAGCACTGCTGGCGCAATCCGTAAGCAGAA 180
Qy 2333 TACATTAAGCGTGGAGATGATGACAGCAGACTATGTTTCCGGGAGCTACATCGCAG 2392
Db 181 TACATTAAGCGTGGAGATGATGACAGCAGACTATGTTTCCGGGAGCTACATCGCAG 240
Qy 2393 CCGCGGAATGGGATGGGTAGAGAGCTATGAGAGCTATGATGATGATGATGATGAT 2452
Db 241 CCGCGGAATGGGATGGGTAGAGAGCTATGAGAGCTATGATGATGATGATGATGAT 300
Qy 2453 CGAGGCGAGTCTGTGTGTGATGAGAGCATTTTGAAGCTCCAGCTACCTAAGGTAATTA 2512
Db 301 CGAGGCGAGTCTGTGTGTGATGAGAGCATTTTGAAGCTCCAGCTACCTAAGGTAATTA 360
Qy 2513 ATCAACTACTCACTTAGCGGAGCACAACCGCTGTGAGCTGAGAGCAGCAGCGGCGCC 2572

```

QY 579 RSISAVDTAAQAKVRAQYTLXSQRKPLGPNNSNOA-----SMTGAAAGSLPRPL 632
Db 1486 RSLP-----TYS--YKPFSTRPYOSWTTAPITVGPAGSGF---TSL 1524
QY 633 NSNSGGTPATP-----GSVTPRP---RGRAGVSQSOPGSRSTSPYKLRDQYGI--- 680
Db 1525 SSSSNTPSASPILKSISVSTPSPILKSTLGASTTSVKSISDVASPIRSLRTWSPKTV 1584
QY 681 -----GNYR--GATGAIKKAGIPRSTASSRETSPTRSGGLMKRSMYSTG 726
Db 1585 VSGSPYNIQVSGTLLAPAPVTEATPLK--GLASNSTFSRSTSPVTTAGLLERSSIT-- 1640
QY 727 AGSRATPERNNPVRPSPAPALLAQSRAEHTLGVDDGQPDYVSGDYMRSGMRMKLM 786
Db 1641 ----MTP---PASPKNIMMYSSSLPFK-----SITSAAPLISPLK 1676
QY 787 G-----RDSDDIDSEASVCSERSPDSYTRGNKSNYSLSGSHRLDWSJORAPPDIE 841
Db 1677 SVVSPYKSRVDYISSAKITWASSLSVPQMPGHAELVNGSISPLKYASS-----S 1729
QY 842 TIIQFCASTHMERKDGILSTLOYLADGKELTQOOLKVLDMRKMFMDHTHTKVISLFD 901
Db 1730 TLINGCKAT-----ATLOEKISSATSVSVSAATDIVEKVFSTTTAMPFSPILRS 1780
QY 902 TYTELLVHANETSRNGSSCLTRLFNKJGTDLNMSHKIMKTLQVHEYPFLOLKE 961
Db 1781 YVS--AAPSPQSLRTPSASA--LTSIGSS--ISAITSVTSITTVPYSVVNVLPPEP 1834
QY 962 LPRILSDSTQPTTKTRIALRFLTDLA-----NTYCKSDPFPDQOACERTVLKLAOL 1016
Db 1835 ALKKLPDUNS--FTKGAALLSPIKTLTETHQPHFRSTSPVKSLSFLAPSAKLTSTP 1892
QY 1017 AADQSMELRSQARSC-----LVALYNLTPOMTLLADLP-----KVY 1055
Db 1893 SSISSSQEILKDVAKEMKEDLMRTAILQTDVPEEKPEOPELPEKGRIDDEEPFKIVEKYK 1952
QY 1056 ODSAR-----SCISHMRQSQSCNGANSPPSSPLSSSPKPLQSPSVGFASLOS 1107
Db 1953 EDLVKQSEILKQDVQDNKSPKSPKSDKG--HSPEDDWIEFSSSEIREA----- 2000
QY 1108 HHHQLSISSTSPRSQSVQELLFSSSELDIQNIQKTEBEIRHCFQGYOTALAPNGFN 1167
Db 2001 --RQQAASQSP--SLPERVQVAKAASEKD--YNLTKVIDYLTNDIGSSSLT----- 2047
QY 1168 GHLQY--HDQGOQDSCASLSNSKTQSSANTQSTPESATMRLDNLERE----- 1215
Db 2048 -NLKYKFEBAKKQBGQKRVLPALALQEHKLMKMP--PASMRTSTSEKEICKXADSFFGT 2105
QY 1216 -----RTQNAKSPPTDAKIVTS---INMAENGELIASNLMESEVVRVALTLT 1262
Db 2106 DTLESPPDDPSQHDQKSPLSDSGFETRSEKTPSAPQSAETTGPKPLFHEVPIPPVITET 2165
QY 1263 KDQPVELLQTSLTNLGICIKGNCELPN-----KHFRS 1295
Db 2166 RTEVVAVIRSYDPSAG-----DVPQOPREBPVSPKPSPTFMELEPKPTSSIKEKYK 2217
QY 1296 IMRMLNIIIEAETDVVIAGLAVLSKIMSNKMRHMMHFLLELILKIIQCYOHSKEALR 1355
Db 2218 AFQMKRASSSEDDHN-----RVLSSKGMRYKEETH-----ITTTVMVHSPPGGE 2261
QY 1356 DIDSMIPRIAPSLPLDLSTINIVPVATGEFPNLCIKILLEVTEHSGSEITDAHLDIV 1415
Db 2262 GASERIEE-----TMSVHDIMKAFQSGRDPG-----KELAGLFHKA- 2300
QY 1416 FPNLASADDTQGMVRAAVFCIVKLYFVLGEEKVXPKLSVLNPSKVRLNLYIEK 1471
Db 2301 SPDVHKSAAETSAQHAKE-----DNQMKPKLE-----RIIEVHIEK 2336

```

```

QY 576 PSMRISAVNTAAQAKAVAOYTLVSRKPKPLGPNNSNOASMTGAASGLRPRILNSN 635
DB 622 PSMNPSSSTPGSSQSTITTEGSAAS-----PTGSGTSTPSVAITEVTSQSTVP---SGS 673
QY 636 SGGCPART-----TGASVTPRPRAGVAGSQSP-----GSRF--TSPTFKLDQYGGIG 681
DB 674 SLGTOSTNSPSSPSSLSPTSGMSTLT--SEPPSPSTOSSGAQSTLTTPSPNPQSTSSLE 732
QY 682 NYVRGATGA-----IPKASGIPRSTASRSTSPTRSGGGLMKRSMYSTGASRRT 732
DB 733 SSTSGATTSAGSAGTWTSPSGSSSVSGSGSTSPASASTTSG-----EMTSQSTGT 784
QY 733 PEKNNPVRBPAPRLAQAQREAEHTLGVGD--GQPDVSGDVYRSGMRKRLMGDRDS 791
DB 785 PGSS-----VSTSAAILSTQOOSVSTNSPGSTVTRPSTVSGSTSGSTVTVG-----S 832
QY 792 DDIDSEASVCSRSPSSSTRGNKSNYSLSGSHTR-----LDMGTORAPPDIE 841
DB 833 TEASTGSSVASSSPADSTQNPVPTSSGSSSMITQSPYSGSTSPVESSTTSPSGPGT 892
QY 842 TTIQFCASHTMSERKQGLISLTD--YLADGKELTQOOLKCYLDPMFKMFMDTHTKYVSLFL 900
DB 893 TL-----TSISPSPSQSTTIGSTGSGTSPGISTSEE-----MTSGSGTQGTGSGTG 938
QY 901 DYTTELILYHANET-----SRNGSSCLTRLFNKLGTDLLNSMHSKIMKTLQYVHXY 952
DB 939 STVTOPSTVSDSTSGSTVTVGSTEGLSSPSP--TSQNTNPSTSSGS-----984
QY 953 FFPQLQKLKFLRISDSTPPTTKTRIALRFLTDLANTYCKSDPDSQACERTVLK 1012
DB 985 -----SMSTQTPQSSQSTSPVESSTSGNTS--SSGSPG-----TTLIS 1020
QY 1013 LQAQADQKSMELRQAQARCLVALY--NLNTPQMTLLADLPKYVQDSARSCIHSMRR 1069
DB 1021 ISSPSPSSSTIGSSQSTSPVVGTSIGSETETGSGTSTVTKPSTVSGASSSGSTATMG 1080
QY 1070 QSGSCNSGANSPPSSP-----LSSSPKPLQSPVVG-----P 1101
DB 1081 TEASSTSGSGSTSPNPSQSTSPSTSGATSSPGSSGTTLTLSISPSQSSSTIGSSQSTSP 1140
QY 1102 FASLQHHHQLGISSTSPRSROSSVEQ-----ELFPSSFLDIQHNQ-----1143
DB 1141 VVSTTGDMTSGQSTQIPGSGTSTVTPSTGSGSTSGEITSGQSTQTPRSSLSTSPAI 1200
QY 1144 KTSEERHRCFGQYOTALAPNGFNHLYQHDQCODSCASLSSNSKT--OSSANTTOSNTP 1202
DB 1201 STSTQOSVSTNSGSTVTPSTVGRGT--SSGSTVTTGSTBSSSTSGSSAITSLSSSSP 1257
QY 1203 ESATMRDLNLERRTTONAKSPDDAKVITVSINMAENGEL 1243
DB 1258 VSTSTQSPN-----PSTSGSSTPTPNPSQSTSPVSVTTTGM 1294

```

RESULT 15
A55575
ankyrin 3, long splice form - human
N:Alternate names: ankyrin G
C:Species: Homo sapiens (man)
C:Date: 03-Mar-1995 #sequence, revision 03-Mar-1995 #text_change 31-Dec-2004
C:Accession: A55575
R:Kordeli, E.; Lambert, S.; Bennett, V.
J. Biol. Chem. 270, 2352-2359, 1995
A:Title: Ankyrin-G. A new ankyrin gene with neural-specific isoforms localized at the ax
A:Reference number: A55575; MID:5138209; PMID:7636469
A:Accession: A55575
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-4377 <KOR>
A:Cross-references: UNIPROT:Q12955; UNIPARC:UP1000013C497; GB:U13616; NID:G608024; PIND:
C:Genetic:
A:Gene: GDB:ANK3
A:Cross-references: GDB:424503; OMIM:600465
A:Map position: 10q21-10q21

```

C:Superfamily: EGF homology
C:Keywords: alternative splicing; peripheral membrane protein
F:73-105/Domain: ankyrin repeat homology <AN01>
F:106-138/Domain: ankyrin repeat homology <AN02>
F:139-171/Domain: ankyrin repeat homology <AN03>
F:172-200/Domain: ankyrin repeat homology <AN04>
F:201-233/Domain: ankyrin repeat homology <AN05>
F:234-266/Domain: ankyrin repeat homology <AN06>
F:267-299/Domain: ankyrin repeat homology <AN07>
F:300-332/Domain: ankyrin repeat homology <AN08>
F:333-365/Domain: ankyrin repeat homology <AN09>
F:366-398/Domain: ankyrin repeat homology <AN10>
F:399-431/Domain: ankyrin repeat homology <AN11>
F:432-464/Domain: ankyrin repeat homology <AN12>
F:465-497/Domain: ankyrin repeat homology <AN13>
F:498-530/Domain: ankyrin repeat homology <AN14>
F:531-563/Domain: ankyrin repeat homology <AN15>
F:564-596/Domain: ankyrin repeat homology <AN16>
F:597-629/Domain: ankyrin repeat homology <AN17>
F:630-662/Domain: ankyrin repeat homology <AN18>
F:663-695/Domain: ankyrin repeat homology <AN19>
F:696-728/Domain: ankyrin repeat homology <AN20>
F:729-761/Domain: ankyrin repeat homology <AN21>
F:762-794/Domain: ankyrin repeat homology <AN22>
F:795-827/Domain: ankyrin repeat homology <AN23>

Query Match 2.5%; Score 192; DB 2; Length 4377;
Best Local Similarity 18.1%; Pred. No. 0.11;
Matches 303; Conservative 242; Mismatches 613; Indels 518; Gaps 76;

QY 69 SLEAFSELKRGSPFNATATVLPVHVIDRLDSDSTREKQQLLRDMHEVRVLPQYL 128
DB 906 SLGARSASLRSRSD--RST-----LNRSYADSMWIB--ELVPSKEQLLFTFREPD 956
QY 129 IDKL-----ATSCFGHK--AKVREBFLOT--IVNAL--HEXGTOOLSRYVYIP----- 172
DB 957 SDSLHYSAADTLDNVNLVSSPIHSGFLVSMVDARSGMGNHGMRIIIPRKCTA 1016
QY 173 ---VCALL-----GDPTVYREAIQTLVEIYHVGDRLEPDLRRMDVP----- 214
DB 1017 PRTICRLVKKRKLAMPNPHGRGRISSRLVEMGPAQFLQPLV---VEIHPGSMRK 1073
QY 215 ASKLML-----EKKPDVQKQGLLPALKNTNNGVGLDEADNIGRERTPMI 265
DB 1074 ERELVLRSNGETWKEQFDS--KNEDL-----TELLNMDELDSPBELG--KKRICR11 1126
QY 266 KR--PLHSAVSSSLRKPKN--VNDVTGDAGAVTMESEFSSF-----EVVPO 307
DB 1127 TQDFQYFVVSRIQESNQIQPEGLISSTVPLVQASFPGALTKIRVGLQAPVP-- 1185
QY 308 LNIFFAKMDIYKQVL-----V1ISDRNADMEKRV----- 338
DB 1186 -----DEIVKILGNKATFSPVIVTEPRRRKFKIPTWIPVPPSGSGVSNQYK 1236
QY 339 DALKKIRALLISYTPQFQFAVQKELS-----LSFV-----DIL 374
DB 1237 DTPNMLRLICSTIGTSP--AQWEDITGTPILFIKDCVSPFTTNVSARFMLADCHOVL 1292
QY 375 KE-ELRSQVIREACTTIAYMSK-----TLRKLDAFCSS--ILEHLINLIQNSAKY 422
DB 1293 ETVGLATQYRE--LICVPMKGFVVPFAKNQDVESLSRFQCMTDKDKVDKTLIQOENFEVY 1351
QY 423 IASASTIALK---YIKYTAAPKLLK-----IYTDTLNQ---SKSKDIRSTLCEL 466
DB 1352 ASKQIEVLEKPIYVDQGNLAPLTKGQQLVFNFSFKENRLPFSSIKIRDTSGPCOR 1411
QY 467 MTLFEEMQTKLENNATV--LQDTL-----KKSIGADQCARHNSR---YAYVAFRRHPE 518
DB 1412 ISFLERKTKKLPQTAVCNNLITLPAHKKETESQDDDEIKTDRQSFASLALRRKRY 1471
QY 519 LADQVGTLDIAQAALBREREGGGGCGTGTGAPETRRIVSRIGRTGLQKPTPM 578
DB 1472 LIE-----PQMIENSTGAT 1485

```

D
760 SGTQSQSVSTSLADSSASTSGSIIVSTSASTKSTSVLSLDSVASKSLSTSESN 819

Y
738 PVPSAPARLLAQSREAEHTLVGDGDDOPDVVSVDYRSGGMWKMGRDESDIDSE 797

D
820 SV-SSSTSTVLNVSQSVSSMS--GSVKSTSLSISNSNSTEKSSEL-STSTDRLRS 876

Y
798 AS-SVCSESPDSSYYTRGNKSNYSLSGSHTRLDMSTORAPPDIETIIQFCASHWERK 856

D
877 TSLSDLSMSTSGSLSKSQSLSTISGSS-----STASLSDSTNALSTSLSESAST 931

Y
857 DGLISLTQYLADKELTQQQLKCYLDMPFKMFMTHTKYVLFELDT-----VTLEILVA 910

D
932 SDPISISINSLANQSASATSK-----SDSQSTISLSTSDSKSMSTSESLED 977

Y
911 ANETSRNGSSCCTRLFNKLGTDLNLSMHKIMTKTLQVWHHYFPQLKLEFRILSDST 970

D
978 STTSAGSVSGSLSLAAQSIVSTSDMSITS-----ELVSDSI 1015

Y
971 QPTTYRTIALIRFLTDLANLYCKSDPFSDQACERTYLKLAQLAADOKSMEL-RSQ 1029

D
1016 STSGSLDS-----ADSCKMSVSSSMSTSGSGSTSLSDSQSTSDSDSKSLSTSGS 1068

Y
1030 RCLVALYNLNPQTMLLADLPKYODSARSCCHSHMRROSGCN-----SGANSPPSS 1084

D
1069 GS-----TSTSTSTA---SVRTSESQSTSGSMSASQSDMSISTSTSDSTSDSKS 1116

Y
1085 PLSSSSPKPLQSPVPFPAFLQSHHHQLSTS-STSPRSQSSVBQELLFSSELDIQHNQ 1143

D
1117 ASTASEESISQASSTSTSGSVST--STSLSTNSESTSTSVSDSTSLSTSEBSISEST 1173

Y
1144 KTSEEIRHCFGGGYOTALPNGFNGHLQYHDQGQDSCASLS---SNSKTQSSANTTOQN 1200

D
1174 STSDSTSEALISAESTSISLSEN---STSDSSQSQAATLSEBLESSTSTSESVS 1230

Y
1201 TESATMRLDNIEREFITTONAKSPDPDAKYIIVSINNAENGELILANLMESEVVAVLT 1260

D
1231 TSESTSLSDSTSESGSTSLSNSTSGASISISTSTISE-----STTFKSESVESTLS 1284

Y
1261 LFKDQPVELLQTSLTN 1276

D
1285 MS-----TSTSLSN 1293

RESULT 14

T34434

hypothetical protein K06A9.1a - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C/Accession: T34434

R/Geisels, C.; Gattung, S.
submitted to the EMBL Data Library, December 1996

A/Description: The sequence of C. elegans cosmid K06A9.

A/Reference number: Z21525

A/Accession: T34434

A>Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-2232 <GET>

A/Cross-References: UNIPROT:O8IFX6, UNIPARC:UPI0000085219, EMBL:U08046, PIDN:AACT0890

A/Experimental source: strain Bristol N2; clone K06A9

C/Genetics:

A/Gene: CESP:K06A9.1a

A/Map position: X

A/Introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 1088/1; 1367/1; 2039/1; 2049/1; 207

Query Match 2.5%; Score 193; DB 2; Length 2232;
Best Local Similarity 21.1%; Pred. No. 0.033;
Matches 173; Conservative 98; Mismatches 334; Indels 216; Gaps 31;

Y
545 GGCTGTGTCT-----APTTRKTVSRIGKTPG--TLQKPT 575

D
568 GSTSTGOSTLASSTATPGSSSTVPSSSPQSSQSGAPMTGTF-----TPSQTSQSPS 621

QY 1076 SGANSFSSPLSSSPKPLQSPVGPASLQSHH-----QUSIS-----STSPRSQSSVE 1127
 DB 4273 SRPESVASQP--ESVSPSPSGAA-----SHEHKEVEISESHKAKSRPESVASQVS 4322
 QY 1128 QELL-----FSELDIQHNIQCTSEIRHCPCGOVOTALAPNGFNHLYOHQGOQDS 1180
 DB 4323 EKDKTRSPASSTQSFSTKEGDEETTSLLH-----SLTTTETVETKOMEKSSPES 4374
 QY 1181 CASLSNSKTQSSANTQSNTPESATWRLDLERERTTONAKSP--TDDAKVTVSINM 1237
 DB 4375 VSTSVTSTYLSLSSSTQY-----LREESTSESLSSSLKXEDSRRRSLSL 4421
 QY 1238 AENGELLASNLME-----SEVVRVALTTKQDPVELLOTSLTNLGICTKGNCEL 1288
 DB 4422 AEKGIGANTSTLKEDTSASASQLEELVQSECESESIVSEIQTSI-----AOK 4470
 QY 1289 PNKFRSIRMRLNLILAEHRTDVYIAGLHVLSKIMRNMNMHLELILKICQYQ 1348
 DB 4471 SNK-----EIKARERTKVTSTQTTTTSATKQDSLKEVIAEF--LATEKIVAKE 4518
 QY 1349 -HSKEALRDISMIPRIAPSLPLDLISINVPVATGEFPTNLCAIKILLEVEHSGSEI 1407
 DB 4519 AFSFEATKRSADDCIKTTASAVSSTASQRLPFGTDBSRRE-----SLSQASB--SRL 4571
 QY 1408 TDAHLDIVFPYLARSADT-----OSMVRKAAVFCIVKLYFVLGEEKVPKLSVLNPS 1460
 DB 4572 T--HSD--PEDEPADVDVDRSSVSKSRKSIATIMWTSTY-----KPSDEMEPIS 4618
 QY 1461 K-VRLNLVYIEKQNCISGGGSSST-----KNSSAASS 1492
 DB 4619 KLVEEHEHVEELAEVSTSTKTTLLQSSSEOSTTSS 4657

RESULT 11
 Probable transposon protein [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C/Accession: F86403
 A:Medeolgie, A.; Ecker, J.R.; Palm, C.J.; Federpriel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chinn, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Hutzar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: F86403
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1148 <STO>
 A:Cross-references: UNIPROT:Q9G6N3, UNIPARC:UPI000009EBB8, GB:AE005172, NID:q10998923; F
 C/Genetics:
 A:Map position: 1

Query Match 2.6%; Score 198; DB 2; Length 1148;
 Best Local Similarity 20.0%; Pred. No. 0.0061;
 Matches 233; Conservative 136; Mismatches 432; Indels 362; Gaps 56;

QY 516 FPELADQ-----IYGTLDIAQALEREEREGGGGCTGTGTATAPET----- 557
 DB 108 FPLSDDPAPASVVRGRPOSQISLSRSTMEKSRSSKGSAPRRLSTSRADNMQOIR 167
 QY 558 -----VSRIGRTPG-----TLQKTPSMESISAVDTPAAQRAKVR 595
 DB 168 GRPSAHPSPASGRSGTTPVRRISPTGKPSGVPVSRPTLSRMSVSTTMSA--- 224
 QY 596 AQTLYSFQRKPLGPNNSNQASMTGAASGLP-----PRLNSNSGCTATTPGAVTP 649
 DB 225 -----VAGTSPVSSSRGNSPSPKIKVQNSNIPGFSLDAPNLRITSLDRPASVYVRSASP 278

QY 650 RPR-GRAGYSQ-----SOPGSRSTSPS-TKLADQY-----GGICNY-----Y 684
 DB 279 ASRNGRDAVSTRRKSVSPASRSVSSSHSHEDRPSQSKGVASSGDLLSLQSI PV 338
 QY 685 RGATGAIIPKASGIP--RSTASGRSTSPRSGGGLMKRSMYSVAGASRTPRNN-----P 738
 DB 339 GGSERAVSRKASLSPSRSTSRSSKLSLSP-----GSAPRRFEALQMHPKHHHSFRP 393
 QY 739 VRPSAPARLQAQREA-----EHTLVGDGQPDVYSGDYNMGGMRKMKLRGD 789
 DB 394 LASSLSTGIVSGKSSVHYHMLRSTATVGSNSSGQVGTGMPDAKGM-----D 444
 QY 790 ESDDDISEASVCSERSFDSSTRG--KNSNYSLSGSHRLDMSTQAPAPDDIETITQFC 847
 DB 445 PVPFQSEVENLAVPDKHEESIAFGVNVNLSNESHSHESFSDLDGMDODVTV--EC 502
 QY 848 ASTHMEBRKDGSLSLQVYADGKELQOQOLKCYLDMPKMFMDTHFKVY--SLPLDTVTE 905
 DB 503 ES-----SANEVSHQ--VPDEN--STHGSIHVGNFLEGV-- 536
 QY 906 LILVHANET-----SRNGSSCLTRLPNKLGTDLNLSHGSKIMWTLQVHEVFPQOLKE 961
 DB 537 -----ALETMEVCGRGSHYCAT-----ATRSEINICECREHGFVETD----- 577
 QY 962 LFRITDSTQTPPTTKTRIALRFLTLANTYCKSSDPSPSQACERTVAKLQADLADQK 1021
 DB 578 -----SPGTNSP-----KLSQTFDEN 594
 QY 1022 SMELRQARSCLVALYNLNPOMTLLADLPKV-----YDSARSC 1062
 DB 595 K-----LTFENLPVID-VLDSLPVYVMEELIETBEKIEQCNDSYEQEYHL 640
 QY 1063 IHSMEROSQSCNSGANSPPSSSPKPLQSPVGPASLQSHHQLSISTSPRSR 1122
 DB 641 YESISIRALEQVQVDMNLNYDQSGSTGCPILSIGTQDTQGLDKKHADVNLISLGRGDV 700
 QY 1123 QSSVEQELLFSSSLDIQHNIOKTSEIRHCPCGOVOTALAPNGFNHLYOHQGOQDS 1182
 DB 701 PLVYKSVSKSPV-IOANNS-----CFTRSYE-----YSRURSI 738
 QY 1183 SLSSNSKTQSSANTQ-----SMTPEASATWRLDLERERTTONAKSPFDDAKVTVS 1234
 DB 739 SLRSSTELTASASSMDYGSIRKGSHTRQSSGSLD-LEHRTVDTNKS-----LSTMS 791
 QY 1235 INNAENGELLASNLMS--EYVRVALTTKQ-----QVELLOTSLTNLGICTI 1281
 DB 792 SSSGMSHTQALNWPEDSFEMCAQMTCTLDETHQESHTEPQNLCKETNVNADFE 851
 QY 1282 KGCNCELPNKHFRSIRMRLNLILAEHTDVYI-----AGLHVLSKIMRNMNMH 1333
 DB 852 SVG-----LVKISANVLGDLEHNPVYVMDCECCNGDNVANTVTSKGTRESIPA 900
 QY 1334 HFPELILIKTI--QC--YQSHKALRDISMIPRIAPSLPLDLISINVPV-----AT 1383
 DB 901 HIRSTSDLGASPTTDDCPFNDSRLQENDVNETPHGLSTTASIEPESSEPELPGCVH 960
 QY 1384 GEFF-----TNL-----CAIKILLEVT--EHHS-----EITDAHLDIV-FPN-----LAR 1421
 DB 961 DELPESRRLNAAVDDGSEKSMYASVDHSSSAPVNEILDESVTLVCPGGEKPERSLTLE 1020
 QY 1422 SADDT-----OSMVRKAAVFCIVKLYFVLGEBK-VKRLSVLNSKVRRLNLYIEKQ 1472
 DB 1021 EATDTLIFCGSIHVDLVYQAATIAMDAKDVPAEBEMLHPTVVLGKSN-----N 1071
 QY 1473 RNCIS-CCGGSSTK--NSSAASS 1492
 DB 1072 RNSYGLGGGTGVAKRSSKAAKAS 1094

RESULT 12
 C84586
 hypothetical protein At2g20200 [imported] - Arabidopsis thaliana

QY 747 LLAOSREAHHTLGVDDGQPDVYSGDYMRSGMGRKLMGRDESDIDSSAVSVCERS 806
 DB 611 LKNKIQKYE-----ILSG-----SSVDLPQAEFLSSNL 640
 QY 807 FDSYTRGNKSNYSLSGSHTRLDWSTORAPFDIETII-OFCASTHMERDGLISLTQY 865
 DB 641 TDALYS-GSSICYSLIFSHSLDLTFQ---YVDIASLSOFLCYVDPSNVGSHFALASF 696
 QY 866 -LAQKELTQOOLKCVLDMFRKM- MDTHYKYSLFDLTVELILVHNETSRNGSSCL 923
 DB 697 PYVSHYDAHKYFPPIVEVLNINISMADHVKVP--FNTNQKRLIH-----GCL 744
 QY 924 TRLNKXGTDLDLNMHSHKIMKTLOVHVEYFPQOLKEFRITISDSTQPTTKTRIALR 983
 DB 745 LML-XEISDTYKLDNLENK-----PFVYT-----DKLRYSSKILAMAKTKLP--- 786
 QY 984 FLTDLANTYCKSPDPSQACERTVLAQLAADQKSMELRSQARSCLVALYN-LNTP 1042
 DB 787 -----SKWMIPLSGLL-----FSLRAHDTFMEDGLDRLNBE 818
 QY 1043 QMTLLADLPKYVDS---ARSCIH-SHMRQSQSCNAGNSPSSS-----PLSSSSPK- 1092
 DB 819 SRTLVAVSMWK--QDADYKSSSTHGHLSKNLPTLNTSSSNSSQTDLVPKCKETKE 876
 QY 1093 -PLQSPVGFPAISQSHHQSLSTSS---PRSSQSVQELFESSLIDQHNIOKTS 1146
 DB 877 TEMQSPLESKEGLSKDTHIESPGSTLEKENEENEGKNPYESNC--SEESLDHNDIQTL 934
 QY 1147 EERHCGGQYOTALAPNGFNHQLQYHDQGDSCASLSNSKTSQASANTQNTBESA- 1205
 DB 935 VNKKETLAQDSBELLQNNALNEKGFENQGLSSAAKYLKDTLDHVSPIINSVSSSP 994
 QY 1206 -----TWRLDNLERERTTQ---NAKSPTDKAVTVSINNAENGELILASLMESV 1254
 DB 995 KDFRTPFKEINERETGFELTSYVALSKKD---INVQTEVDESVGIANAFMDN 1049
 QY 1255 VARALLTLTDQPV---ELDQSLTNIGCTKGCNCELPNKHFRSINRMMLNITIEAHTDV 1311
 DB 1050 VNQDSINSYDQSSGKDKLLTSTST-----PNKPTTFMPANEIILGSPAKDY 1096
 QY 1312 VIAGLHVLSKIMSNKRNH 1331
 DB 1097 DHDQSYSHLSENNREN 1116
 RESULT 10
 113564
 microbule-associated protein homolog - fruit fly (Drosophila melanogaster)
 N.Alternate names: hypothetical protein EG:49E4.1
 C.Species: Drosophila melanogaster
 C.Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 31-Dec-2004
 C.Accession: T11564
 R.Spanos, L.; Papadogiannakis, G.; Siden-Kiamos, I.; Louis, C.
 submitted to the EMBL Data Library, April 1999
 A.Description: Sequencing the distal X chromosome of Drosophila melanogaster.
 A.Reference number: Z17689
 A.Accession: T11564
 A.Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-5327 <SPA>
 A.Cross-references: UNIPROT:O76891; UNIPARC:UPI0000110149; EMBL:AL031128; PIDN:CAA20006.
 A.Cross-references: FlyBase:FBgn0025392
 A.Intons: 24/2; 52/3; 104/3; 179/1; 232/1; 1669/3; 2566/1; 4798/3; 5272/1
 A.Note: EG:49E4.1
 Query Match 2.6%; Score 198.5; DB 2; Length 5327;
 Best Local Similarity 19.3%; Pred. No. 0.064;
 Matches 308; Conservative 227; Mismatches 635; Indels 429; Gaps 70;
 QY 102 SRDYREKAQQLLRLDMLHEHRLVLPQALIDKLATSCFKKKNKAVREPIQITVINALHEKGT 161
 DB 3280 SRDSVAEKSPLASKASR-----PASVASVQDEAEKESRRESRRESVAEKSPLAVKENS- 3333

QY 162 QQLSVRYIIPVICALGDTFVNVRBAIQTLVEIYKHVGDRRLPDLRRMDVPASKLAML 221
 DB 3334 -----RPASV---AESIDAEKESKESRRESVAEKSPLASK 3367
 QY 222 EQKFDQVKGEGILLPSALKNTNGVGLDEADNIGRERPTR---MIKRLHSAVSSSLR 278
 DB 3368 E-----ASRPTSAESVQDEAEK--SKESRSRQSVAEKSPASKENS--R 3408
 QY 279 PKRNADVQDAGAVMSEFESSFVYVPLNITHAD-----MDIYKQVLVI 327
 DB 3409 PASVASVQDEA--EKSKESRRESVAEKSPLASKASRPASVASVQD----- 3456
 QY 328 SDKNADMEKRVDAKKIRALLILSYHTQPOFVAVQKELSLSPVDLILKEELRQVIREAC 387
 DB 3457 AEKSKESRRESVAEK--SPLASKESRSPASVASVQDEA---EKSKESRRESVAE-- 3508
 QY 388 ITAYVSKTLRNKLDACWISLEHLINLIONS-----AKVIASASTIALKYIITYHAPK 442
 DB 3509 -----KSPLPKESASRPTSAESVQDEAEKESRRESVAEKSPLASKESRSPASVAE 3562
 QY 443 LKITYDTLQKSKDIRETLCMLVLFEEWQYALERNATYLRDTLKKSIGDADCDAR 502
 DB 3563 SVK--DEAEKESRRESVAEKSPL-----ASKEASRPASV---AESVQDEAEKSK 3609
 QY 503 RSRVAVYAFRRHFP-----ELADQYGLDIAQALAREREGC---GG 544
 DB 3610 EYSRRESVAEKSPLPSKESRPTSAESVQDEADSKSKE---ESRREGKESKPLASMEAS 3666
 QY 545 GGTGTGTGAPTRRTVSRIGRTPTLOKTPSRMSISAVDTAAQAKVRAQYTLYSRQ 604
 DB 3667 RPTSAESVQDETEKESRRESVTEKSPLSKSE--ASRPTSAESVQDEAEKESKEER 3724
 QY 605 RK-----PLGNMNSQ-ASM-----TGAAASGSLPPRLNSN 635
 DB 3725 RESVAEKSPPLASKESRSPASVASIKDEAGTKQSRRESMEPSGKASIKGDQSLASK 3784
 QY 636 SGGTPTATPGSY---TPRPRGAVGVSOGPSRSTPSTKLRDQYGGIGNYRGATGAI 692
 DB 3785 ETSRSDSVESVKDTEKPEGSA-IDKQVASRPSPASVAKDEKSPLSHRPESVADKSP 3843
 QY 693 KKAAGIPTS-----TASSRETSPTRSGGLMKRSMYSTAGASRRTPERNP----- 738
 DB 3844 DASKEASRSLVAETASSPIEEGPRSIADL-SLPNLTGEAKGKLPITLSPIDVAEGPFL 3902
 QY 739 -VRPAPARALLAQSRSEAHHTLGVDDGQPDVYSGDYMRSGMGRKLMGRDESDIDISE 797
 DB 3903 EVKAESSPRPAVLSPAEFS-----QPD-----TGHYASPPVD-E 3936
 QY 798 ASSVCSERSFDSYTRGNKSNYSLSG-----SHTRLDWSTORAPFDIETI---- 843
 DB 3937 ASPVLEIEVEGQHT--TSGVGATGATAETDLDLTETKSEITYTKQSETTLFETLSKV 3993
 QY 844 -----IOFCASTHMERDGLISLTQYLDGKELTQOOLKCVLD----- 882
 DB 3994 ESKVEVLSSVQYVEKQVQTSVQAETTVDSLQGLTK--KSSQLT--EIKSVLDNTNIS 4049
 QY 883 -----MFKMMDHTIKYKYSFLDVTVELILVH-----ANETSRNGSSCLT 924
 DB 4050 NVTNLFSTAVETIEKKVQDVTEK-----IEKATHEVSEHVTTGESSTETSOEKSSLDL- 4104
 QY 925 RLPNKLGTDLDLNMHSHKIMKTLOVHVEYFPQOLKE-----LFPHISST 970
 DB 4105 GTFELMETHHTTYSSEF-TVITCEDEPVLHDIKEDEBHRSPSPSDVUKAIIIPPQ 4163
 QY 971 QPTTKTRIALRLFDLTLANTYCKSDPS-----DOSQACER-----TVLKLQAQLAAD 1019
 DB 4164 MRPLSPREEEVAKIVADAVAKLSKDXDITDIIPDFDRQLEKUKSTADTDEEDSKSTRD 4223
 QY 1020 QKSWEILSQARSCIVALYN-LNTPQW---TLILADLPKYVQDSARSCISHMRQSQSCN 1075
 DB 4224 EKSLIET---SVKVEIESKSSPDQKSGPISIEKDKIBQSEKAQUL-----ROGILAS 4272

RESULT 8

S28293

hypothetical protein ZC94.3 - *Caenorhabditis elegans*C/Species: *Caenorhabditis elegans*

C/Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Jul-2004

C/Accession: S28293

R/Thomas, K.

submitted to the EMBL Data Library, December 1992

A/Reference number: S28285

A/Accession: S28293

A/Molecule type: DNA

A/Residues: 1-643 <THO>

A/Cross-references: UNIPROT:Q03609; UNIPARC:UPI0000178CEA; EMBL:Z19157

C/Genetics:

A/Intons: 141/2; 167/3; 245/3; 341/2; 412/3; 600/3; 640/3

Query Match

Best Local Similarity 3.4%; Score 262.5; DB 2; Length 643;

Matches 140; Conservative 137; Mismatches 265; Indels 175; Gaps 28;

```

QY 796 SEASSVSESRFSDSYR-----GKSNVSLSGSHRLD-----829
DB 40 TRTSITSDSRDTSPTRRNSPLPEPQKARVKYNGSFFAKLGMPDPTDDEFLPIRI 99
QY 830 WSTORAP----FDIETIIQFCASTHSEKRGDLSITQYLADGKELTQOOLKCVLDMFR 885
DB 100 RSLPKTPTIEAHDKVQVLKCCSSSVSEKKEGIKLPIYAD-TSLNPIEIKNIGCN 158
QY 886 KMFENDTHT---KYSLFLDTVELLIVANETSRNGSSGCTLLENKLGTDLINSMSK 941
DB 159 RLISDAENTWLEIYSIFVFT-----HSSRLS-EWLRLALAKLFARKAETLPNTKKQ 210
QY 942 IMKTLQVHEFFPQOLQKELFRISDSTOPTTKTATILRFLTDLANTYCKSSDPSD 1001
DB 211 IGHLLNVLIECFNAHDLVTCMLCPHLMVFKARVLELYTSLIDEYTEBG-ASI 268
QY 1002 QSQACERTVLKLAQLADQK-SMELRSQARSCLVALYNLNTPQMTLLADLPKYODSAR 1060
DB 269 NAKELTKAIRKMLTWADPRLSILLTPHVEKALICMVCNVADFSALISD-----DSEOK 324
QY 1061 SCISHNRROSQSCNGANSPSSPLSSSPKPLQSPVGFPSLQSHHQLSISSTG- 1118
DB 325 NMIHQTQIR-----NGLENGISNNIATNS-----GATASRETSNTSFOKESTFG 370
QY 1119 -PR--SQSSVEQELLSSSELDIQNTQKTESEIRHCFGGGYOTALAPNGNGLQYHDQ 1175
DB 371 LPEFGARKGCTGNL-----GSLINISNNLALSRLE-----400
QY 1176 GQOQSCASLSNSKTSQSSANTQSTNPESATMRLDNERRTTONAKSPTDKAVITVS 1235
DB 401 --EOSTRLMEKVNLANSTV-TLPDPTLEKIONVODLQKMSSENABE--QESALISSIM 455
QY 1236 NMAENG-----ELILASNLMESEVVALTLYTDQPVELLQSTLTNIGICIKGNC 1286
DB 456 MICDGGGVMEQCYAKILL-NLFE-----ILSKSR-----SENNKMC-----492
QY 1287 ELPKHRSIRMRMLNILEAHTDVTAGLHVLISKIRSNKMRNNMHFLELLILKLIQC 1346
DB 493 -----LRIGKCKTAQ-----AAKLFDST-----EMAVCKVLDA 521
QY 1347 YHSEKALR--DIDSMIPRIAPSLPLDLISINIVNPVIATGEPFN-----LCAIKILLEVT 1400
DB 522 AVNTNDATLALAVEDCLRTLTHPLPSNIIIAKYL--NOEPLDDEBASLVKQVTRLF 579
QY 1401 EHHGSEITDAHLDIVFPMLASADDTOSMKRAAVFCIKLYFVLGEEKVKPKLSVL 1457
DB 580 EELPAEELNINIVDITPTIIKAYOSTSSTVAKTVVYCLVAMVNRVGEQRMTPHPTKL 636

```

RESULT 9
T11648probable mitotic spindle protein - fission yeast (*Schizosaccharomyces pombe*)C/Species: *Schizosaccharomyces pombe*

C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C/Accession: T11648

R/Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, December 1997

A/Reference number: Z17304

A/Accession: T11648

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1462 <MOO>

A/Cross-references: UNIPROT:Q42874; UNIPARC:UPI000069DB2; EMBL:AL021046; NID:e12167f

C/Genetics:

A/Map position: 1L

A/Intons: 17/1; 39/3

A/Note: SPAC39.12

Query Match

Best Local Similarity 3.4%; Score 259.5; DB 2; Length 1462;

Matches 272; Conservative 211; Mismatches 564; Indels 353; Gaps 57;

```

QY 1 MAYRKPSDLIDGFIQMPKADMRVKVQLAEDLVTELSDDTNSIVCTDMGFLIDGLMPVLTG 60
DB 1 MADKQADDFLKLKSNASTD--EKTRCLDTLRSEF--NKNININADIGLFECEFRLLATT 56
QY 61 SHFKIAOKSLFANSEILKRLGSDFNAYTATVLPH-----VIDRQDSRDVYREKAQLLRD 116
DB 57 VNPFLRLSSVACFETFLRRRAQYPTWLKFRVYMLKNVLVDHIA-SRDLOKRVNLITL 114
QY 117 LMEHRYLPQALIDKLATSCFKHNAKVREEFQITV---NALHEXTQOLSVRVVYIPV 173
DB 115 LMFNPESEIEKSLIHIST---SKSAETRIQCKMVLAINNALSDPVKSLRALYIN-- 169
QY 174 CALLGDPYVNVREAAIQTVIEIKYHVGDRLPDLRRMDVPASKLAMEOKFPQVQOGL 233
DB 170 ---LENANPSVREAEKVELLIYKMLSTSAK--MQFTIVDETTSGLRREILQSLVEBELSL 224
QY 234 LPLSA---LKNTNGCVGLDEADNIGLERPFRMKRPLHSAVSSSLRKPKNVNDVTGD 289
DB 225 ISSSSEVIIVQNS-----ASSFQAPFMTAVATL 253
QY 290 AGAVTWSEFSSFEVYVPOINIFPAKMDIDYKQVLVIISDKNAD--WEKVDALKIRAL 347
DB 254 YPQVELEEN-----VYPLANF-SKQLEODSASMLPAFERERFQKMSVQSDVLRURQY 306
QY 348 L-----ILSYHTQOPFAVQAKELSLSFVDILKEELSQVIREACITTAIVNSKTLRNKLA 403
DB 307 LRGNACIDY--LPELLISV-LKTL-LPGIILLALSLRTTLSSSAIQLIKEMAIILKSNIDP 362
QY 404 FCWSIIEHLINLIQNSAKVIASASTIA-----LKTIITYTAPKLLKIYTD 449
DB 363 FLELIIPNLKVCYVTKLASQANVTFAAILVNCGLSRNLSFISLAH-----D 413
QY 450 TLMQSKQIRSTLCMLVTLFPEWQTKA-LEBNATVLRDTLKKISIGDADCDARHGRYA 508
DB 414 T--NAQLRFPSSWIMFLLISLSPELKXLASLQNLKAFELICRGGLADSSQVREYVRKS 471
QY 509 YNAFRHFPFLADQIYGLDIAAQRALEREREGGGGGGTGTGTAPETRRVSRIGRTP 568
DB 472 FMTLSEYFPEVQSELNLTLEPSVLKQHL-----499
QY 569 GTLQKPTPSMRSISAVDTAAQAQAKYRAQYTL--YSRQRPGLPNNNSQMSMGAAASGS 626
DB 500 --LANPN--RQAASFNFGPKAPRIPSLNLSFSGSQEETSSNSNSSGT-----547
QY 627 LRPPLNNSGCPATTPGVSVPTRPRGRAGVSOQSGSRSTSPSTKLRDQYGGIGNYRG 686
DB 548 ---RRLGLPQKATPAGRE-RVLPYTRSQAHSLSLPSLSGHSPT-----590
QY 687 ATGAIPKKAAGIPRSTASRSTSPTRSGGLMKRSMYSTGAGRRRTPERNNPVPSAPAR 746
DB 591 ---AIPS-----KRSVATIKDSKTFE-----610

```



```

Db      107 AATCSPTVSKYGIETHSIGEDILVPMASOVAVKTKMTASTLTFEIVEYQTRVFT 166
      446 IYDTLLNOSKSDIRSTLCELMVLLFEEMQTKALERNATVLRDLTKKSIDADCDARRHS 505
      167 ILS-SFSTSDKSRRLALALLEIVISKMSDRIKQIMROI CELIKSAINDADESETRPAAG 225
      506 RYAWAFRRHPPELADQIYGTLDIAQPALEREEREGGGGTGTGTAPETRRYTRIG 565
      226 RRAFAKIDEMHSEADALYLELHDSKQKMLR----- 256
      566 RTGCTLOKPTPSMKSISAVDTAAQAKVAQYTLYSRQKRPGLGNNSNQASMTGAASG 625
      257 ----- 256
      626 SLPRRLNNSGCTPATPGSVTPRPRGRAGVSQSQSGRSTSPSTKLADQYGGIGNYR 685
      257 -----GDAASSMASV----- 267
      686 GATGAIIPKASGIPRSTASRETSPTSGGGLMKRSMYSTGAGSRRTPERNNPVPSAPA 745
      268 -----NSEKSGSIP 276
      746 RLAAQSRBAHTLGVDDGDPDYVSGDYMSGGMKRMKMGDESDIDSEASVCSER 805
      277 R----- 277
      806 SPDSYTRGNKSNYSLSGSHRLDMSQORAPFDDETIIQFCASHTWSEKDLISLTQY 865
      278 ---SKSAGSKAHNNISAI-----SEKKEGKILLP 306
      866 LADGKELTQOOLKCVLDMFRKFMNDTHT---KYVSLFDLTVELLIIVHANEYSRNGSSS 921
      307 VAD-TSLNPIEIKNGICNLRLSDASNTVLEIYSIFVRI-----HSSRLS-EWLRL 357
      922 CLTRLENKGLTDILNMSHAKIWKTLQVHEFPQLOLKEFLIISDTGTPTTKRTIAI 981
      358 ALAKLFARKAETLPNTRKQIGHTLVILECFNHHDLVYCELMCDPIHLMVKAIVL 417
      982 LRFLTDANTYCKSDPSPDOSACERTVLKLAOLADOK-SMELRQASCSVALNLN 1040
      418 LEYITSLIDEYTEBG--ASINAKELKTAIRKMLTWASDPRISILLTHVEKALISMCVN 475
      1041 TPQWTLILLADLPKYVQDSARSQSHMRROSOCSNGANSPPSSPLSSSPKPLQSPSVG 1100
      476 VADSALISDL---DSEQKWIHQIOTR-----NGLENGISSNNIATNS-----G 517
      1101 PPSALQSHHQLSISSTS---PR--SROSSVEQELSSSELDIQHNIQKTSSEIRHCFGG 1155
      518 ATASRETSNTSFOKESTSFGLPBFGARKGTGVNL---GSLINISNNLALSRL----- 567
      1156 QYQALALPNGFNHGLQYHDQGOQSCASLSNSKTOSSANTQSPESATMRDNLERE 1215
      568 -----EOSTSRHLEKNVNLNTV-TLPDPTLEKIONODLLOK 604
      1216 RTTONAKSPTDAAKIVTWSINMAENG-----ELLASNLMESEVVRVALLTYDQ 1266
      605 RSSENADE--QESAISSITMMICGGFGVMEQCAKILL-NLEF-----ILSKSR- 651
      1267 VELLQTVLNLGICIRKGNCELPNKHFRSIRKMLNLLEAHTVYAGLHVLSKIRSN 1366
      652 -----SENNKMC-----LRILGKCKTQO-----AAKLDPST 678
      1327 KMRHNMHFEILLIKIIOCYHSEKALR--DIDSMIPRIAPSLPLDSINIIVPVIATG 1384
      679 -----EWAICKVLDAAVNTNDATTAALAVEDCLRTLATHPLPSINIINIAVIL--N 726
      1385 EFPNTN---LCAIKILLVEVTEHSGSEITDAHLDIVFPNLARSADDTQSMVKAAVFCIV 1440
      727 QEPIDDERASIVLKMVTRLPELPAFELNNIVDDITPTIIKAYOSTSIVKTVVYGLVA 786
      1441 LYLVLGEEKVKPKLVNPSKVRLNVYIEK 1471

```

```

Db      787 MNVRVGEQRMTPHFTKLPRKAMTNLIQVYVNR 817

RESULT 6
S68176
TOG protein - human
C/Species: Homo sapiens (man)
C/Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C/Accession: S68176
Eur. J. Biochem. 234, 406-413, 1995
R/Characterization of the cDNA and pattern of expression of a new gene over-ex
A/Reference number: S68176; MUID:96128167; PMID:8536682
A/Accession: S68176
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1972 <CHA>
A/Cross-references: UNIPROT:Q14008; UNIPARC:UPI000016787C; EMBL:X92474; NID:91045056;

Query Match      4.3%; Score 328; DB 2; Length 1972;
Best Local Similarity 18.7%; Pred. No. 7, 2e-10;
Matches 278; Conservative 253; Mismatches 543; Indels 416; Gaps 68;

      61 SHEKIAQSLAEFSELIKRGLSDPNAYTATVLPVVIDRIGDSRDIVREKAQLLRDLMEH 120
      653 TNEQVQMKLHIVA-LIAQKGNFSTSAQVLDGLVDKIDYK--CGNNAKEMTAIAEA 709
      121 RVLPQALIDKLATSCFKHKNKVRBEFLQITVNALHEVGTQQLSVRYIIPYCALLGDP 180
      710 CMLPMTA--EQVSNMFSQKPNQSETLNLMSNAIKERFGSGLVKAFISNKTALAA 767
      181 TNNVRAAIQTLVEIKVHVDRLRPLRRM--DDVASTLAMEOKFDDQKQGLLPSA 238
      768 NPAYRTAATLLGVMVLVYG---PSLRMFEEDEKA-LISQIDAEF--KMGQSGPPAP 820
      239 LKNTNGNV-GLDEADNIGLRERPTMIKRLPSHSAVSSSLRPKNVNDVTDGAAVTMS 297
      821 TRGISGHSSTGDEGD-----GDEPDGSDNV----- 849
      298 FESSFEVNPQLNIFKAKMDDIYQVLIISDNKADWEKVDALKIRALLISYHTPOQ 357
      850 ---DLRPTETL---SDKITSELVSKIGDN--WKIRKEGLDEVAIGI-----NDAK 892
      358 FVAVLKEISSLVVDLKELELSQVREACTITAMSKTLRNGLDFMCSIEHLNLIQ 417
      893 FIQPNIGELPTA---LKGRLNDS-----NKI-----LVQQLNLTLQ 925
      418 NSAKVIASASTALKYIIVYTHAPKLLKIYDTLLNOSKSDIRSTLCELMVLLFEEMQTK 477
      926 Q---LAVAMGNPKQHVKNLGIIP-IITVLGDSKNVRAAALATV----- 965
      478 ALERNATVLRDLTKKSIDADCDARRHSRYAWAFRRHPPELADQIYGTLDIAQPALER 537
      966 ---NMAAGTQKEMLEBEDLSEB-----LKKNPFLAQEELGWL---AEK--- 1005
      538 EREGGGGGGTGTGTAPETRTVSRIGRTPTGLQKPTSMRSISAVDTAAOR---AKV 594
      1006 ---LPTLRST-----PTDLILCVPHLYSGLERNGDVVK 1036
      595 RAQYTL-----YSRQRKPLG--PNNNSQASMTGAASGSLPRRLNNSGCTPAT 643
      1037 KADALPFPMMHGLYKMAATGKLKPTSDQ-----VLAMEKAKVMPPAKPAP--- 1086
      644 PGSVTPRPRGRAGVSQSQSGRST--SESTKLADQYGGIGNYRGATGAIIPKASGIP 699
      1087 PTKATSKPMGGASPAKFPASABEDCISSTSPKPD-----PKA----- 1127
      700 RSTASRETSPTSGGGLMKRSMYSTGAGSRRTPERNNPVPSAPARLLAQSREAEHTLG 759
      1128 ---KADGLSSKAKSAQG---KKMPSKTS-----LKED----- 1154
      760 VDDGQDPDYV---SGDYMSGGMKRMKLM-----GRDE-SDDIDSEASVCS---ER 805

```

QY 1445 LGEEKYKPKLSTVINSKRVLLNVIIEK 1471
 DB 1027 LGKMTMEPHLONISGKLNLVQVYVNR 1053

RESULT 4

330876
 hypothetical protein R107.6 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995
 C:Accession: S30876

R:Thomas, K.
 submitted to the EMBL Data Library, July 1992

A:Reference number: S30871

A:Accession: S30876

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1080 <THO>

A:Cross-References: UNIPARC:UPI000017BB14; EMBL:Z14092

C:Genetics:

A:introns: 74/3; 342/1; 560/3; 688/3; 816/3; 963/3

Query Match 7.8%; Score 593.5; DB 2; Length 1080;
 Best Local Similarity 20.7%; Pred. No. 3.4e-25;
 Matches 261; Conservative 208; Mismatches 472; Indels 319; Gaps 39;

QY 288 GDAGAVTMESFESFEVYVQNLNIFPAKMDIYKQVLVVISDNKDMEKRVALKKIRAL 347
 DB 54 GGTCTVSKDDFLKSFEEVYVQNLNIFPAKMDIYKQVLVVISDNKDMEKRVALKKIRAL 113

QY 348 LILSYHTPOFVAVQKELSLFVDILK--ELRSQVIREACITIAVMSKTLRNLDAF 404
 DB 114 VV---HGEDVIGREQLSQLVRLTDCLDLSVKDLRSQILREAAITCGFLFKFGTDVQRI 170

QY 405 CMSLIEHLINLQNAKVIASASTALAKIITKATAPKLIKTYTTLNOSKQDRLSLC 464
 DB 171 AERCPSPAPQAVATKVMATGAVLTFTVEFIQTKFCIA--SYSTSKDKNRRLC 229

QY 465 ELMVLLPEEMOTKALERNATVLDTLKKSIGDADCDARRSHRYAVAFRRHPELADQY 524
 DB 230 ALLEVLVLEHNMKIRKTVLPQIGELIKALICDADETRVAGKAKSKLDALHSTADLKF 289

QY 525 GTLDIAAPALEREEREGGGGGTGTGTAP-ETRTVSRIG--RTPTLQKPTPSMSI 581
 DB 290 ASVDSKQKMLRASDAASSTINSERGTAPFRSKLSAGSIGIRNAPNISKFLAQRSA 349

QY 582 SAVDTAAAOBARKVROVTLYSQRKPLGPNNSNQASMTGAASGLPRRLNSNGSTPA 641
 DB 350 SAIDTKQVTRMAT---SVSRTP---NIRPMTRTILS---KIDTSPGSGKF 390

QY 642 TTP--GSVTPRP---RGRAGVOSQPGSRSTSPSTKRDQYGGIGNYRGATGAIIPKA 695
 DB 391 ARPVTGALGSTRSSNLRRAGSVPTSPGSRNSPPRR-----PSATEFPPEM 438

QY 696 SGIPRSTASSETSPSTRSGGGLMKRSMYSTGAGSRRTERRNN-----PVPSAPARLLA 749
 DB 439 QRVKSNLNSNFSVSLSAEATKLOKAMTAKESLRQPSRNDDEFLPKKPT--PQKATP 497

QY 750 QSRREHTLVGDDQDPYVSGDYRSGGMKGRKLMGRDSDIDLSASVCSERSFDS 809
 DB 498 QK----- 499

QY 810 SYTRGNKSNYSLSGSHRLDWSSTORAFDDIETIIQPCASTHSEPKDGLISTLOYLADG 869
 DB 500 -----SALDTSR-----VEEVIKACSSSTSNKREKIGKLAGIVSE 535

QY 870 KELTQOOLKCYLDNERKPMFMDTKVYSLFDVTTELIVANETSRNGSSCITRLFNK 929
 DB 536 PNLNAEIKSIGAVLNRLIGESTNQI---VLESISSFVKTTHPRLS--DMLKLGKGLPAK 591

QY 930 LGTDLNMSHKIMKTLQVNHVYPTQOLKELFIISDSOTPTTKRIALLFLFDL 989

DB 592 KGAENTLNSKKQISTTISLSSFPDITQLKSTCELVCDPIHLMSPKSRVLLLEYNELL 651

QY 990 NTYC-KSSDPSPDSQACERTVLKLAQLAQKMEI-RSQARSLVAYNLNTPOMTLL 1047
 DB 652 GKIMERGSSFTYKMKMA---TILKMFMMADQRRNEQLITPGEKVLCSLFALNADFFAL 708

QY 1048 LADLPKYQDSARSCISHMRRQ-----SQCNSGANSPPSSPLS-----SSSPK 1092
 DB 709 FNDPNPDYRDWAVYKVLQSHGHQHVPOODAVSEACVATISTTAAQIEDEVRNLDMT 768

QY 1093 PLOSPVSPFPASLSHHQSLSSSTSPRSRSSVEQL--FSSSELDIOHNIQKTSER 1150
 DB 769 FVKSPT------RAISSGPKK-----VDAPPLRPLSEKMSQHRDELS---- 807

QY 1151 HCFGQYOTAPALPNGFNQHYHDQGDSCASLSNKSQTQSANTQSNTPESATMRD 1210
 DB 808 -----FN-----ESTDRKLNSTHLLIDTSE----- 829

QY 1211 NLBERRTONAKSPDQDAKIVTYSINM-AENGELILASNLMESEVRVALTLYDQVEL 1269
 DB 830 -----QSKYVASKLAQISGDMGAQOYEGILS----- 855

QY 1270 LQSTLTNLGICIKGNCNELPKKPRSRIMRLNLLEHMDV-VIAGLHYSKIMRSKYM 1328
 DB 856 IQTMLCE-----GSFTLEQNFALKLIAVFDVLSKSESPANKKVALRVLTQCTSOAS 908

QY 1329 RHNMMHLELLILKIIQCYOHSKEALRDI--DSMIPRIAPSLPLDSINI----- 1376
 DB 909 R--LFDSTELALICVLDAAANSQDGTNNVATDCLKTLATLPLAKYVINSQLLNEKA 966

QY 1377 -----VNPVIANGEPFTNI--CAIKLLEVTEHHSSEITDAH 1411
 DB 967 QEPKASIVLKMMTRELFEGQADELSPVV-----DDLAPCVIKYITVQTSQFINEIF-- 1017

QY 1412 LDIVFPLASADPTQSMWRKAFCIVKLYFVIGEEKVYKSVLNPSPKRVLLNVIIEK 1471
 DB 1018 -----TLQSYSPSSAVKRTAYVCLVAVNKKLMKMTMEPHLONISGKLNLVQVYVNR 1070

RESULT 5

C88550
 protein ZC84.3 [imported] - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: C88550

R:anonymous, The C. elegans Sequencing Consortium.

Science 283, 2012-2018, 1998

A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating bio

A:Reference number: A75000; MUID:90659613; PMID:9851916

A>Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.e

A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; a

A:Accession: C88550

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-826 <STO>

A:Cross-References: UNIPROT:Q03609; UNIPARC:UPI000013BAAC; GB:chr_III; PIDN:CAA79568.1

A:Gene: ZC84.3

A:Map position: 3

Query Match 5.1%; Score 389; DB 2; Length 826;
 Best Local Similarity 17.1%; Pred. No. 6.8e-14;
 Matches 211; Conservative 203; Mismatches 371; Indels 446; Gaps 35;

QY 273 VSSSLRKPVMNVDTGAGVMTESFESFEVYVQNLNIFPAKMDIYKQVLVVISQNA 332
 DB 1 MTSRMDPKSGVY-----SVSKADFTKIFEDVYKVPITSAVDLRNKFDAVRITLSNSSE 53

QY 333 DWEKRVDAKKIRALLILSYH-----TOPOFVA--VOLKELSLFVDILKEEBSQVIRE 385
 DB 54 DMNRQGTQKTVASLVI---HGEKVVDPRPTMIAHLQV-----LGCEFLAVKDLRSQVIRE 106

QY 386 ACTTIYMGKTLNKKLDAPFWSLIEHLINLQNSAKVIVASASTIALKYIITKTHAPKLLK 445

```

QY 930 LGTDLNSMSKIMKTLQVNHVEPTQLOMKEFLIISDTPPTTKRIALIRFETDLA 989
: : : : :
Db 775 MGADLLSVAKQVAKALDVTRRESFNDQFNILMFYDQITQTPSKVAVILIKTYETLA 834
: : : : :
QY 990 NTFCSDPSFSD--QSOACERTVLKLAQIADQKSEMRSQARSCLVALYNINTPOMTLL 1047
: : : : :
Db 835 ----KQMD-PGDFINSETRLAVGRVITWTTEPKSDVRKKAQSVLISLFELNTPETVL 889
: : : : :
QY 1048 LADLPKYVDOSARSCIHSHMRROSOCSNGANSPSSPLSSSPK-----PLOSPSVG 1100
: : : : :
Db 890 LGALPPTFOGATYLLNHNLH-----NTNGTOSMSGSPLTRPTPRS PANMSSPLTSPVT 945
: : : : :
QY 1101 PRASLOSHHQLOLSSSTSPRROSSVE-----QLLSPSELDIQNIQKTSB---INH 1151
: : : : :
Db 946 SQNTLPSPAEDYDTENNNSEDIYSSLRGVTAIQFSPRSQDMMPELRDSKKDDGDSM 1005
: : : : :
QY 1152 CFQGOYOTALAPNGFNGLQYHDQGOQDSCASISNSKTOSGANTQSN---TPESATM 1207
: : : : :
Db 1006 CGS----PGMSDPAGGDAITSSQOTALDNKASLSHSMPTHSSPRRDYVPPYNSDSISPF 1061
: : : : :
QY 1208 RLDNLERERTQNAKSPFLDPAKVITVSINMAENGELILASINMESEBVAVALLTYQDPY 1267
: : : : :
Db 1062 NKSALKEAMFDDDDADQFPDD-----LSLDHSD--LVAELLKEKS-----NHNERV 1104
: : : : :
QY 1268 ELLQJVLTNIGTICKGNCCELPNKGFRSINMNLNILEAHRTDVVYAGLHVLSKTIKRSUK 1327
: : : : :
Db 1105 EERKIALYELMKLTQEBSPFSWDEHFKTILLLETTDQKEPTTALAKVIREILRHOP 1164
: : : : :
QY 1328 MRHNMMHFLLELILKIQCYOH--SKEALRDISMTPIRLAPSLPLDISINVPVIAITGEF 1386
: : : : :
Db 1165 AR--FKNYABLYWMTKEAHNDHPKEVVRSAEASVLATISISPGCIVLCPITQIOTAY 1222
: : : : :
QY 1387 PTNLCAIKILLEVEHHGSEITDAHLDIVPNLARSADDTOSMRKANVFCVLYFVLG 1446
: : : : :
Db 1223 PIVLAAIKQOTKVIERYSKETLANELLPEIMGLLIQYDNSEBSYKCAVFCIVAAVAVIG 1282
: : : : :
QY 1447 EEKVAKRLSVLNPBKVAVLLNVIIEKORNCISGGGSTKXNSAAS 1490
: : : : :
Db 1283 DE-LKPHLSQLTGSKMLNLNLYI--KRAQTGGGADPTTVDVSSQS 1324
: : : : :

```

```

RESULT 3
E88546
protein R107.6 [imported] - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: E88546
R/Anonymous: The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A/Reference number: A75000; PMID:99069613; PMID:9851916
A/Note: see webistes genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A/Note: published extrata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A/Accession: E88546
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1063 <STO>
C/Cross-references: UNIPROT:P32744, UNIPARC:UPI00000610FC, GB:chr_III, PIDN:CAA78472.1,
A/Gene: R107.6
A/Map position: 3

```

```

Query Match      8.0%; Score 611; DB 2; Length 1063;
Beet Local Similarity 20.9%; Pred.No.3,4e-26;
Matches 257; Conservative 213; Mismatches 487; Indels 270; Gaps 36

QY      288 GDAGAVTMSFESSFEVVPOLNIFPAKDMDDIYKCVLIVISDNADMEKRVADLKIRAL 347
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       54 GGTCTVSKDPLKSFEEVPKMEISPSDFEKLDQTIETLSGQEDMNKRNKKLKQIISM 113
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      348 LLTSHYTOPFAVAVQKELSLSFVDLKK--EELNSQVIREACITIAVMSKTLRNKLDAF 404
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db 114 VV---HGEDVIGRRQLLSOLVRLTRLDCCDLSVOKLRSQILAEAAITCGFLFRKRPSTVDROJ 170
Qy 405 CWSILEHLINLIONSANVASASTIALKVIITKHAPBLKITYDTDLNOSKODIRSTLC 464
Db 171 AERCLPSAPFAQVAVASTKMAATCGAVLFLIFIEIOTKOIFETCIA-SYSTSDKNOROLC 229
Qy 465 ELAVLLEPEWOTKALERNATVLRTLKKSIGDDDCDARRSRAYMAFRHFPBLAOIY 524
Db 230 ALLEIVIEHNNEKIKRTVLPOIGELIAAICDADPEFRVAGRKAFSKDALHSTEADKLF 289
Qy 525 GTLIIAORALEREREGGGGGGTGTGTAP-ETRRVSRIG--RTGTLOKPPMSRSI 581
Db 290 ASVSSSKOKMLASDAASSSTISNSEGTAPEFRKLSAGSIGGRNAPNISSKRLAORSA 349
Qy 582 SAVDTAAAOARAKVRAOYTLVSRORKEPLGPNNNSQASMTGAAAGSLRPLRNSNGTGA 641
Db 350 SAIDTKQVTRNAT-----SVSRTP-----NLRPTTRTILS-----KIDTSPGSKF 390
Qy 642 TTP--GSVTPRP---RGRAVGSOSQSGSSTSPSTKLRPOYGGIGNYGATAIIRKA 695
Db 391 ARPVVALGSRSTSNLRFARGSVPTSPQSRNGSPRR-----PSATEAFPAEM 438
Qy 696 SGIRSRSTASSHENSPTSGGGLMRSMVSRGAGRRTPERRNN-----PVRSAPEARLA 749
Db 439 QRVSNISGNSFVSSLSAEATYKIQKAMNTRAKESLRQPSRNDDEFLPKRPT-POKATP 497
Qy 750 QSRAEHTLGVGDGQDPYVSGDYMRSGMRGKLMGRDESDDIDSEASSVCSERSFDS 809
Db 498 QK-----499
Qy 810 SYTRGNKSNNSLSGSHRLDLSOTRAPPDIEITIQFCASTHMERDGLISLTOYIADG 869
Db 500 -----SALDTSR-----VEEVIRACSSSTANKEKEGIKMLAGIYSE- 535
Qy 870 KELTQOQOKCVLDMFRKMFMDHTHKVYSLFDPYTELILVHANTSRNGSSCCTRLEFK 929
Db 536 PNLSENAIEKSIQAVLNLRLGESTNOI---VLESISSTVKTHHPLS-DMLKLGIGKIFAK 591
Qy 930 LGTDLNLSMSHSKIWKTQVNEHYEPTLOLKELEFRIISDSTOPTTKIRIALNPLTDLA 989
Db 592 KGAEMLTNSKKQJSTTISCLISSDPLLOLQKSTCELVCDPHLMSPSRVLLLEYMELL 651
Qy 990 NTYC-KSSDPSDOSQACERTVYKLAQIADQSMEL-RSQASCVALYNLNPOMTLL 1047
Db 652 GKMYERGSSFFTKEMKA---TILKFMWMDQREBOLITPGEKVLCSLPALNNADPSAL 708
Qy 1048 LADLPKYQDSARCSHSHKRO-----SOSCSGANSPESSPLS-----SSSPK 10922
Db 709 FNDNRPDRIWAAVYVLVLSHGHDQVPOODAVSEACVRATITSTAAQIEDFVSRINDMT 768
Qy 1093 PLOSPSVQFPASLOSHHOLISSTSPRSROSVEOEOL--FSSELDIOHNIQKTSSEIR 1150
Db 769 PVKPEST-----RAISSGPR-----VDAPPLRPLSENNSSQHRDELS--- 807
Qy 1151 HCFGGQYOTALAPNGFNGHLQYHDGQODSCASLSSNSKTQSSANTTQSTPESATMRDL 1210
Db 808 -----FN-----ESPDRKLNSTHILIDTSE----- 829
Qy 1211 NLEBERTQNAKSTPDAKVIITVAINM-AENGELILASNLMESEVAVALLTMDQPEVL 1269
Db 830 -----QSTRYVASKLAQISGDMAQOYEBELS----- 855
Qy 1270 LOTSILNIGICIKGANCELPKHFPRSIMRMLNLTLEAHTDV-VIAGLHVLSKIMRSNMK 1328
Db 856 IOTMLCE-----GSTPLMEQNRAKLIIAFVLDVLSSESDANKKVALRVLTWCTQAS 908
Qy 1329 RHNMMHFLLELILKTIQCYOHSKEALRDI--DSMIPIABSLPILDSINTVNPVI--ATG 1384
Db 909 R-LFDSTEMAICKVLDAAVNSQGTWNVATADDCIKTLIATHLPLAKVNIISQILNEEKA 966
Qy 1385 EEPFLNCAIKILLEVEHSGSETDAHLDIIVPVLASADDTQSMKAAVFCVXKLYFV 1444
Db 967 QEPASIVLKMNTLIFGLQADELSPVVDLAPCVISYDSSPSAVKRTAYCVLAAVNNK 1026

```

LENGTH: 154
 TYPE: DNA
 ORGANISM: Drosophila melanogaster
 US-09-270-767-30401

Query Match 2.9%; Score 149.4; DB 3; Length 154;
 Best Local Similarity 99.3%; Pred. No. 4.9e-34;
 Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2154 GGAATATCCCGAAGCAGACCCAGCTCTAGAGAAACAGTCAACTAGAGTCAAGTGTGTC 2213
 DB 1 GGAATATCCCGAAGCAGACCCAGCTCTAGAGAAACAGTCAACTAGAGTGTGTC 60
 QY 2214 TTGATGAAGCAGATATGTAATCTCTACAGTGTGCGGTCTCCAGTACCGCCGAGAAAC 2273
 DB 61 TTGATGAAGCAGATATGTAATCTCTACAGTGTGCGGTCTCCAGTACCGCCGAGAAAC 120
 QY 2274 AACCAGTAAAGCAGATCGGCGCGGACGAC 2304
 DB 121 AACCAGTAAAGCAGATCGGCGCGGACGAC 151

RESULT 8
 US-10-104-047-517
 Sequence 517, Application US/10104047
 Patent No. 6943241
 GENERAL INFORMATION:
 APPLICANT: HELIX RESEARCH INSTITUTE
 TITLE OF INVENTION: NO. 6943241el full length cdna
 FILE REFERENCE: H1-A0105
 CURRENT APPLICATION NUMBER: US/10/104,047
 PRIOR FILING DATE: 2002-03-25
 PRIOR APPLICATION NUMBER:
 PRIOR FILING DATE:
 NUMBER OF SEQ ID NOS: 4096
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 517
 LENGTH: 2524
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-104-047-517

Query Match 2.8%; Score 141.8; DB 3; Length 2524;
 Best Local Similarity 46.1%; Pred. No. 6.2e-31;
 Matches 769; Conservative 0; Mismatches 852; Indels 45; Gaps 7;

QY 77 GAAAGCCAGGACCTGATGCTTATTCAGCAAAATGCCAAGGCGAATGCGTGA 136
 DB 357 GAGAGCCCGCAGCATGAGTACTTCTGCGCCAGGTGACAGAGACGTGCGGCG 416
 QY 137 GGTACAGCTTCCGAGGATCTGTGACATTCCTTAGCG---ACGACAAATCAATTG 193
 DB 417 GCTGAGGTGCGCCAGAGCTCTGCTCTACCTTGGGCCCCCGGCCCACTTCGACCT 476
 QY 194 CTGACAGCAATGAGTCTCT-----TATTGAGGTTTGTGTCATGCTGACGG 244
 DB 477 GAGAGGAGACCTGGCGCGCTAGGCAAGACAGTGCAGCGCTCACCGGCTGAGTTC 536
 QY 245 CAGCCACTTAAGATTGCAAAAGTCCCTGAGGCGCTTCTCGAGCTATTAAGCAT 304
 DB 537 GAGCACTACCGGATCATTAATGAGATGGAATTTAAGTGCCTTGTGACAGATT 596
 QY 305 GGGCAGCATTTTAATGATACACGGCTACCGTTCTGCCACATGTGATCGATCGGCTGG 364
 DB 597 ATCAACACGCTTAAATCTATGATGCAATGTTATGTTATGCTTAAATAGACAGATGG 656
 QY 365 AGACAGAGGAGACAGCTCCGCGAGAGGCGCAACTTCTCTCGCGGACCTCATGAGCA 424
 DB 657 AGATGCCAAGAGAGGTTGAGATGAGTGAAGCTCAGACTCTGATATTGAAGTTAATGATCA 716
 QY 425 CAGAGTGTTCGCGCCAGGCGGTGATGACAGCTGCGCACTAGCTGCTCAAGACAA 484
 DB 717 AGTATG-----CAGCAGCTATGATCAATTGGAGCGAGTGGCTTCTGTTTAAACAA 770

QY 485 GAAAGCCAGGTCGCGAGAGTCTCTCAAGCAATGTGAAAGCGCTTCATGAGTACGG 544
 DB 771 GAAATTTGATCTGAGAGAGGCGTGTCTGTGTCTTATGTAACCTTAACATTTTGG 830
 QY 545 CAGCCAGAGCTTATGTTGCGGTCTATATACCAACCACTTTGTGACATTCGAGATCC 604
 DB 831 GGTCTAGGCACTATGATCAACCAATGATACCAATTTGTGTATCTGTTTGAAGACTC 890
 QY 605 CACAGTTAATGTAGGAGGCGGCATCAACGCTATGTGAAATCTCAAGCATGTAG 664
 DB 891 CACAGTATGATGAGAGTGTGCAATATTTGCTATATGTGAGATTATTAAGCATGTGG 950
 QY 665 GATTCGATTTGGCCAGACCTCCGTGATGAGAGATTTCTCCG-----TCGAAATTGG 720
 DB 951 AGAAATGATGAGATGATCTTTATTAAGAGAGATTTCCCTCTGTATATTAAGATGAT 1010
 QY 721 CTATGTTGAGCAAAAGTTGACCAAGTCAACAGAGGCTCTACTGCTTACCTTCAGCC 780
 DB 1011 ATTTCGCAAAATTGATGAGTCAAAAGTTTCAGCGGTATGATTTGAGTGTCTGCAAGA 1070
 QY 781 TTAATAACAGAAATGGAATGAGTGGCTTGGACGAGCGCAATATTTGGTTGAGG 840
 DB 1071 TAAAGCTTGATGATGAGATGATGAGATGAGAAATAGGCAATCATCATGATCAGC 1130
 QY 841 AGCAGCCACAGGATGATTAAGCGGCCACTACACTCGGCCGTTTCG-----887
 DB 1131 CTTCAGATTCCTGACACTTAACATCCGAAATCTCCCAAGTGCAGAGAAAGCTTGG 1190
 QY 888 -TCATCACTCGCCCAAAACCAGATGTGAACGATGACCGGTAGCCGCGCGTAAAC 946
 DB 1191 TTCAGCAGTGGCCCTTAAGGTGGAAGTGTCTTCAAGAAAGAGTGTGAGCAGTTGA 1250
 QY 947 CATGATCTTTGGAATCTAGTGTGAGTGTGTCCTCCCAATGAACTTTCACGCTAA 1006
 DB 1251 TGAAGATATTTATTAAGACCTTTTACAGATGTCCTTATTCAGATTTATTTAGTCG 1310
 QY 1007 GCAATGAGCAGATATCTACAGCAAGTATGATCATGATGATTAAGCAAGCAGACTG 1066
 DB 1311 AGAATCGAAGAAACATTAATTAATTAAGCAAGAAATTTGTCAATATTAACATGACTG 1370
 QY 1067 GGAAGAACTGTGATGCTCTCAAGAAATGAGGCAATGCTCATTTCTAGCTATCACAC 1126
 DB 1371 GATATAGCGCGCAATGACATGAAGAAATTCATCATCTCTGTT-----GCTGAGC 1424
 QY 1127 TCAGCGCAGTTGTGCTGTATACGCTAAAGAAATGTGCTTAACTTCTGTCGATCT 1186
 DB 1425 TGCACAGTATGATGCTTTTTCACATTTACGATTTGTGATGAGACCTTAAACTTTC 1484
 QY 1187 CAGAGAGAACTACAGATCAAGGTATCCGCGAGGCGTGATCACCATGCGCTTACATGTC 1246
 DB 1485 AGCTAAGATCTTAATCCAGGTGTGAGAGAGCTTGTATTAATCTGTAACCCACTTTC 1544
 QY 1247 TAAAGCTGAGAAATTAACATGATGCTTCTGCGAGCAATTTGAGACCTGATTA 1306
 DB 1545 AACAGTTTGGGAAACAAAGTTGATCATGCGCTGAAGCATGTACCTTAACTTTTAA 1604
 QY 1307 TTTAATACAGAACAGGCGCAAGTCAATGATCCGTTCCACATATGCTTGAAGTAT 1366
 DB 1605 TCTCGTCCCAATAGTCAAAAGTCAATGCAACTTCTGATGTGACGACATCAGATTAT 1664
 QY 1367 CATTAAGTATACATGACCAAGGTGCTCAAGATCTACAGACACTGTAATCATC 1426
 DB 1665 CATTCGCAATCATATGTAACCAAGCTTATACCTTTAATTAACAGCAATTTGCA---CATC 1721
 QY 1427 AAAGTGAAGACATPAAGTTCACACTGTGAGCTGATGTGTCTCTTCAGAGAGTG 1486
 DB 1722 AAATCAGTTCCTCGAGAGAGAGCTTCAATTTAATTTAAGTTATATGTTGAAGAGTG 1781
 QY 1487 GAGAGCAAGGCTTGAAGAAAGATGCCACCTTCAAGGACACTTAAATAATTCAT 1546
 DB 1782 GAGACTCATTCATTGAGAAAGCATGACAGCGCTTGTGTTGAATATTAATAAGGAAT 1841


```

Db      422 ATTGAGATTATTATGTCGAGAGACTCGAAGAAACATTAAATCAAGGAAATTTTG 481
QY      1047 AGTATATAAAACGAGACTGGAGAAACGTGTGATGCTCTGAGAAAGATCAGGCGATTG 1106
Db      482 TCAGATGATTAACATGACTGGGATCAGCGTGCATATGCACTGAAGAAATTTGATCAGT 541
QY      1107 CTCATTCGAGTATCAGACTAGCGCGATTGTTCGTCTGACCTTAAGAAATTTGTCG 1166
Db      542 CTGTGTCTGTGAG-----CTGCACAGATATGATTCCTTTTCAACATTATGATGATGTTG 595
QY      1167 TTAAGCTTCGTGACATCCTCAAGAGAGAACTACATGATCAGATGATCCGAGGCGTGC 1226
Db      596 GATGAGAGACTTAATCTTACGCTAAGATCTTAATATCCAGGTGTGAGAAAGCTTGT 655
QY      1227 ATCAGCATGCGCTACATGCTTAAGACGCTGAGAAATTAATAGATGCTTGTCTGAGC 1286
Db      656 ATTACTGTAGCCCACTTTCAACAGTTTGGAAACAGTTTGAATCATGAGCGCTGAAGCC 715
QY      1287 ATTTGGAGCACTGATTAATTTAATACAGAAACGCGGAGAGTTCATTCGCTTCC 1346
Db      716 ATTGACTTACACTTTTAAATCTGCTCCCAATAGTGCMAAAGTCATGGCAACTTGTGA 775
QY      1347 ACAATAGCTCTGAGTATATCATTTAATATACATGACGACCAAGCTGCTCAAGATCTAC 1406
Db      776 TGTGACAGATCAATTTATATCTGCTGCGATCTCATGTAACCAAGCTTATACCTTTAATA 835
QY      1407 ACAGACACTCTGAATCAATCAAGTCGAAAGACATAGTCCACATGCTGTGAGCTGATG 1466
Db      836 ACAAGCAATTGCA---CATCAAAATCAAGTCCCGTAGAGAGACGTTCAATTTGAATTTTA 892
QY      1467 GTGTGCTCTTCGAGAGAGTGCACAGAGGCGTGTGAAGAAATGCCACCTGACTAAG 1526
Db      893 GATTATATGTTGCAAGAGTGCACATCTATTCATTTGAAGACATGACGCGTCTTGTGTT 952
QY      1527 GACACCTTAAAAAATCCATTTGGGATGCGAGACCTGCGATGCAAGCCCGCATTCAGATAC 1586
Db      953 GAAACTATTAAGAAAGGAAATTCATGATGCTGACCTGAGCGCAAGTGGAGGCAAGAAAG 1012
QY      1587 GCCTATTTGGGCTTTCAGGCGTCACTTTCCAGAGCTGGCGGATCAAAATATATGAAACTTA 1646
Db      1013 ACATACATGGGTCTTAAGAAACCACTTTCCTGTGAGAGTGAATATATATATATTCCTT 1072
QY      1647 GACATAGCTCCCGCGCGCTTGAAGAAAGGACGAGAGGCGCGGAGAGAGAACT 1706
Db      1073 GAGCCACTTATACAGAAAGTCTTCAACTTAAGAGTTCTGGCAGTGTAGCATCT 1132
QY      1707 GTTAC 1711
Db      1133 CTTCC 1137

```

RESULT 11

US-09-620-312D-323

Sequence 323, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

```

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jien-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Dermanac, Radoje T.

```

```

; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725.
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: PL_Fl_genes Version 1.0
; SEQ ID NO 323
; LENGTH: 6487
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (164)..(4093)
US-09-620-312D-323

```

```

Query Match      2.3%; Score 116.6; DB 3; Length 6487;
Best Local Similarity 48.7%; Pred. No. 4e-23;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

```

```

QY      927 GGTGATGCGCGCGCGTAAACCATGAAATCTTTGAAATCTAGCTTTGAGGTGCGCCAA 986
Db      362 GGAGGTGCTGGAGCAGATGATGAAGATTTTATAAAGCTTTTACAGATGTCCTTCT 421
QY      987 TTGAATCTTCCACGCTAAGACATGAGCATATCTACAGCAAGTACTAGTATCATC 1046
Db      422 ATTGAGTTTATTTCTAGTGCAGAACTCGAAGAAACATTAATTAATCAAGGAAATTTTG 481
QY      1047 AGGATATAAAACGACACTGGAGAAACGCTGATGCTCTCAAGAAATCAGGCAATG 1106
Db      482 TCAGATGATTAACATGATGCTGATGATGATGATTTTATAAAGCTTTTACAGATGTCCTTCT 541
QY      1107 CTCATTTCACTATACACTGACGCGCACTTTGTGCTGTACGCTTAAGAAATTTGTCG 1166
Db      542 CTGTGTGCTGAG-----CTGCACAGTATGATTTGCTTTTTCACATTTACATTTGTTG 595
QY      1167 TTAAGCTTCGTGACATCTCTCAAGAGAGACTACATCAGATGATCCGAGGCGTGC 1226
Db      596 GATGAGCACTTAATCACTTTCAAGCTGATGATCTTAAGATCCAGGTGTGAGAAAGCTTGT 655
QY      1227 ATCAGCATGCGCTACATGCTTAAGACGCTGAGAAATTAATAGATGCTTGTCTGAGC 1286
Db      656 ATTACTGTAGCCCACTTTCAACAGTTTGGAAACAGTTTATATATGATGCGCTGAAGCC 715
QY      1287 ATTTGGAGCACTGATTAATTTAATATACAGAAACGCGGAGAGTCAATGATCCGCTTCC 1346
Db      716 ATTGACTTACACTTTTAAATCTGCTCCCAATAGTGCMAAAGTCAAGGCAACTTGTGA 775
QY      1347 ACAATAGCTCTGAGATATCATTTAATATACATGACCAAGCTGCTCAAGATCTAC 1406
Db      776 TGTGACAGATCAATGATTTATCAATTCGCAATCTCATGATCCCAAGCTTATACCTTTAATA 835
QY      1407 ACAGACACTCTGATCAATCAAGTGAAGAGACATPAAGGTCCACATGATGATGAGTGTG 1466
Db      836 ACAAGCAATTGCA---CATCAAAATCAAGTCCCGTAGAGAGACTTATTAATTTTA 892
QY      1467 GTGTGCTCTTCGAGAGAGTGCACAGAGGCGTGTGAAGAAATGCCACCTGACTAAG 1526
Db      893 GATTATATGTTGCAAGAGTGCAGACTCTTCAATTTGAAGAAAGCATGACAGCCGCTTGTGTT 952
QY      1527 GACACTTAAAAAATCCATTTGGCGATGACAGCTGATGACAGCGCCCATTTCCAGATAC 1586
Db      953 GAAACTATTAAGAAAGGAAATTCATGATGCTGAGCGTGAAGCCAGTGAAGCAAGAAAG 1012
QY      1587 GCCTATTTGGGCTTTCAGGCGTCACTTTCCAGAGCTGGCGGATCAATATATATGAAACTTA 1646
Db      1013 ACATACATGGGTCTTAAGAAACCACTTTCCTGTGAGGCTGAAACATTTATATATTCCTT 1072
QY      1647 GACATAGCTGCCAAGCGCGCTTGAAGAAAGGACGAGGCGCGGAGAGAGAACT 1706

```

Db 1073 GAGCATTATATCAGAGAGCTTCAACTTACTTAAGAGCTTCGAGAGTAGCATCT 1132
QY 1707 GGTAC 1711
Db 1133 CTTC 1137

RESULT 12

US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHIEFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match 0.9%, Score 47.2; DB 2; Length 7218;

Best Local Similarity 6.5%; Pred. No. 0.023; Mismatches 89; Indels 0; Gaps 0;
Matches 16; Conservative 141; Mismatches 89; Indels 0; Gaps 0;

QY 4865 CTCGCAATCTCACTTCTCATCATGCTAGCTCTTTTACGAGCTCCTAATAA 4924
Db 1218 YY 1277
QY 4925 CGCGATTCGTATTAATTGTAAGCCTATACCGCTCTTAATGAACACGAGTG 4984
Db 1278 YY 1337
QY 4985 CTTTCTACCTTTGTTGAGGAGCTGTATCTTAATTGTTCCGCTCACTAAT 5044
Db 1338 YY 1397

QY 5045 CTAAATCTATTTATTAATAGATTCGCTCTTACAACTGACATTTTGTAAATTT 5104
Db 1398 YY 1457
QY 5105 AAATAC 5110
Db 1458 AACTAC 1463

RESULT 13

US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHIEFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match 0.9%; Score 43.8; DB 2; Length 7218;

Best Local Similarity 8.9%; Pred. No. 0.24; Mismatches 177; Indels 0; Gaps 0;
Matches 36; Conservative 190; Mismatches 177; Indels 0; Gaps 0;

QY 1337 ATCCGCTTCACATAGCTCTGAATATATTAATGATACATGACCAAGCTGCT 1396
Db 1517 ATTCAGTTTCAAAAAGGAGATGATGACATCTGAATTAATCTATGCAAGTAGTT 1458
QY 1397 CAGATCTACACGACACTGTGAATCAATCAAGTGAAGAGACATAGCTCACTG 1456
Db 1457 AAAGAGATGAAGAAATTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1398
QY 1457 TGAGTGATGAGTCTCTCTGAGGAGTGACGAGAGGCTTGAAGAGATCCAC 1516
Db 1397 RRR 1338

QY 1517 CGTACTAAGGACACCTTAAAAATCCATTGGCGATGACGCTCCGATGCACGCCGCCA 1576
 Db 1337 RRR 1278
 QY 1577 TTCGATAGCGCTTATGGCTTTTCAGGGCTCCTTCCAGGCTGGCGGATCAATATA 1636
 Db 1277 RRR 1218
 QY 1637 TGGACATTAAGCATAGCTGCCCGCCGCAATTAGAAAGGACGAGCGCGCGAG 1696
 Db 1217 RRR 1158
 QY 1697 AGGAGAACTGCTACTGGGACTGGGACTGCACCTGAAAGCA 1739
 Db 1157 RRR 1115

RESULT 14

5244792-1/c
 Patent No. 5244792
 APPLICANT: BURKE, RAE L.; PACHL, CAROL; VALENZUELA, PABLO D.T.
 TITLE OF INVENTION: EXPRESSION OF RECOMBINANT GLYOPROTEIN
 B FROM HERPES SIMPLEX VIRUS
 NUMBER OF SEQUENCES: 19
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/587,179
 FILING DATE: 20-SEP-1990
 PRIOR APPLICATION NUMBER: 921,730
 FILING DATE: 20-OCT-1986
 APPLICATION NUMBER: 597,784
 FILING DATE: 06-APR-1984
 SEQ ID NO: 1:
 LENGTH: 3472
 5244792-1

Query Match 0.8%; Score 42.8; DB 9; Length 3472;
 Best Local Similarity 54.4%; Pred. No. 0.29;
 Matches 86; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 3068 CCAGAGCCAGGCTCCGAGCGGAGCGTCTTAAAGCTGCCCGCGGATCGAA 3127
 Db 1508 CGAGAGGAGTACTGGGTGAGGTGGTGAAGTGTGAGATGCGTCCGAGAGAA 1449
 QY 3128 GTGATGAGCTGGCTCCAGGCGGAGCTGCTAGTGGCCCTGTATACTGAATAC 3187
 Db 1448 GCGAAGAGACCGCCGCTACTCGGCGGAGCATCTGTCACCTCTGCACTTGTCAT 1389
 QY 3188 CCGCAATGACCTTTTACTGGCGGACCTGCCAAGG 3225
 Db 1388 GGTGACAGACCGCGGTGCTTGGCACCAAGTCCAGG 1351

RESULT 15

US-09-561-077C-25
 Sequence 25, Application US/09561077C
 Patent No. 6706501
 GENERAL INFORMATION:
 APPLICANT: Rosson, Reinhardt D.
 APPLICANT: Deng, Ming-de
 APPLICANT: Grund, Alan D.
 TITLE OF INVENTION: LINOLEATE ISOMERASE
 FILE REFERENCE: 3161-20-C1
 CURRENT APPLICATION NUMBER: US/09/561,077C
 PRIOR FILING DATE: 2000-04-28
 PRIOR APPLICATION NUMBER: 60/141,798
 NUMBER OF SEQ ID NOS: 80
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 25
 LENGTH: 3684
 TYPE: DNA

ORGANISM: Lactobacillus reuteri
 US-09-561-077C-25

Query Match 0.8%; Score 41.6; DB 3; Length 3684;
 Best Local Similarity 50.0%; Pred. No. 0.68;
 Matches 130; Conservative 0; Mismatches 129; Indels 1; Gaps 1;

QY 4744 CAAATGGCGCCGAAAGAAAGTCAAGGCGCATCCGCCATCTAATGTGACAGAAAACCTA 4803
 Db 3342 CATTGGGGGAGATTTTAAGGCTATTGGATGACATCTACTAAGCAACAAACATTA 3401
 QY 4804 TTTATACATAACGGGAACCTAAACCGGTAATCGAAGTACTTAAATTTAGCCA 4863
 Db 3402 TTGAACCTAGAAAGCAAGTAGAAGAAATTAATTAATTAATTAATTAATTAATCA 3461
 QY 4864 ACTCGCAATCCTACTCTCACATCGATTGATGCTTTTACAGCTCCTAATTA 4923
 Db 3462 GCGAAGAAATGACCTTACCGGAGCGGAATACCGCTTTGGCAAGCCAAATTAATTTTA 3521
 QY 4924 ACGCTGATCTGTTTAAATTTGTAAGCCTATATACCGCTTTTATGAAACTCAGACGTGT 4983
 Db 3522 TTAATTAATGTG-TAAACTGACAGGSCACAGTACCTGTTTGAAGGAGCTCACAAGCT 3580
 QY 4984 GCTTTTCTACTTTGTTTCA 5003
 Db 3581 GCTTTTCTACTTTGTTTCA 3600

Search completed: November 30, 2005, 00:50:42
 Job time : 594 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 23, 2005, 14:49:57 ; Search time 211 Seconds
(without alignments)
3106.887 Million cell updates/sec

Title: US-10-030-850-2
Perfect score: 7611
Sequence: 1 MAYRKPSDLDGFICQMPKAD.....RNCISGGSGSTKSSAASSS 1492

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 244163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 244163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	7611	100.0	1492	4	AAB31597	Aab31597 Amino aci
2	1707	22.4	1526	3	AAB29634	Aab29634 Human pol
3	1352.5	17.8	1298	7	ADF60378	Adf60378 Human con
4	1348	17.7	1324	7	ADB79873	Adb79873 Rat putat
5	1348	17.7	1324	9	ADX07258	Adx07258 Cyclin-de
6	1344.5	17.7	1295	4	AAM39287	Aam39287 Human pol
7	1344.5	17.7	1309	4	AAM39286	Aam39286 Human pol
8	1344.5	17.7	1317	4	AAM41072	Aam41072 Human pol
9	1344.5	17.7	1317	4	AAM41073	Aam41073 Human pol
10	1214.5	16.0	916	7	ADB64143	Adb64143 Human pro
11	920	12.1	723	8	ADU20651	Adu20651 A. thal
12	608.5	8.0	1330	8	ADB63929	Adb63929 Human pro
13	502.5	6.6	454	7	ADB63929	Adb63929 Human pro
14	497.5	6.5	432	5	ABB81781	Abb81781 Proteasom
15	328	4.3	1972	9	ADJ68457	Adj68457 Human hea
16	328	4.3	1972	9	ADJ68457	Adj68457 Human hea
17	314.5	4.1	2038	6	ABR83341	AbR83341 Human LBD
18	304.5	4.0	2040	6	ABR83341	AbR83341 Human LBD
19	281.5	3.7	306	7	ADJ70616	Adj70616 Human hea
20	276	3.6	407	8	ADJ12968	Adj12968 Plant ful
21	232.5	3.1	2018	4	ABB71336	Abb71336 Drosophil
22	215	2.8	1684	2	ADH11336	Adh11336 Vertebrat
23	215	2.8	2340	2	ADH11419	Adh11419 Human hu
24	215	2.8	2365	7	ADD93763	Add93763 Human hel

25	215	2.8	2429	7	ADD93765	Add93765 Human hel
26	214	2.8	2429	9	ADY98003	Ady98003 Nav2 prot
27	210.5	2.8	2432	3	AAV85565	Aav85565 Human hom
28	205.5	2.7	1831	6	ABU43109	Abu43109 Protein e
29	205.5	2.7	1870	6	ABU19019	Abu19019 Pathogen
30	204.5	2.7	2344	4	AAU37120	Aau37120 Staphyloc
31	201.5	2.6	2261	6	ABJ18914	Abj18914 Pathogen
32	201.5	2.6	2271	6	ABM72734	Abm72734 Staphyloc
33	196.5	2.6	2283	6	ABP56876	Abp56876 Staphyloc
34	196.5	2.6	2271	6	ABU20428	Abu20428 A. thal
35	195	2.6	2271	6	ABU16000	Abu16000 Protein e
36	194.5	2.6	2681	6	ABJ19025	Abj19025 Pathogen
37	193.5	2.5	2137	5	ABP39618	Abp39618 Staphyloc
38	193.5	2.5	2137	8	ABSO5656	Abso5656 Staphyloc
39	193.5	2.5	3201	4	ABR62899	AbR62899 Drosophil
40	192	2.5	1150	4	ABR59129	AbR59129 Drosophil
41	192	2.5	4274	4	ABG00972	Abg00972 Novel hum
42	192	2.5	4397	4	ABG21944	Abg21944 Novel hum
43	190.5	2.5	2426	2	ADH11285	Adh11285 Vertebrat
44	187	2.5	1079	7	ADJ71179	Adj71179 Human hea
45	187	2.5	4386	4	ABG07375	Abg07375 Novel hum

ALIGNMENTS

RESULT 1	AAB31597	standard; protein; 1492 AA.
ID	AAB31597	
AC	AAB31597;	
DT	30-APR-2001	(first entry)
DE	Amino acid sequence of a Drosophila polypeptide designated orbit.	
XX	Orbit; chromosome segregation; mitotic spindle; microtubule; mitosis;	
XX	polyploid cell; microtubule binding; nucleation; motor protein; cancer;	
XX	leukaemia cell; solid tumour.	
OS	Drosophila sp.	
EH	Key	Location/Qualifiers
FT	Domain	290..1093
FT	Domain	/note="HR1 domain"
FT	Domain	326..342
FT	Domain	/note="putative GTP binding domain"
FT	Domain	400..403
FT	Domain	/note="putative GTP binding domain"
FT	Domain	544..549
FT	Domain	/note="putative GTP binding domain"
FT	Domain	1093..1271
FT	Domain	/note="HR2 domain"
PN	WO200104295-A1.	
PD	18-JAN-2001.	
PF	11-JUL-2000; 2000WO-GB002662.	
PR	13-JUL-1999; 99GB-00016402.	
PR	24-DEC-1999; 99GB-00030707.	
XX	(UYDU-) UNIV DUNDEE.	
XX	Avides MDC, Deak P, Glover DM;	
XX	WPI; 2001-138345/14.	
XX	N-PSDB; AAF24991.	
XX	Novel Drosophila orbit protein and its human homolog, useful for treating	
PT	tumors and in assays for identifying substances capable of inhibiting	
PT	mitosis.	

Claim 5; Fig 7; 92pp; English.

The present sequence represents *Drosophila* polypeptide designated orbit. The orbit protein is required for chromosome segregation. Orbit protein localises to mitotic spindles and binds microtubules. Orbit mutant phenotypes result in embryos exhibiting abnormal mitosis and polyploid cells. Orbit polypeptides are useful for inhibiting mitosis, and for identifying substances which affect orbit functions such as microtubule binding, microtubule organizing centre, nucleation activity and interactions with microtubule motor proteins. Tests for detecting or sequencing orbit, or its homologue, in a biological sample may be used to determine orbit sequences within cells in individuals who have or are suspected to have, an altered orbit gene sequence, for example within cancer cells including leukaemia cells and solid tumours such as breast, ovary, lung, colon etc

Sequence 1492 AA:

Query Match 100.0%; Score 7611; DB 4; Length 1492;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAYRRPSLDGFIQMPKADRVKQVLAEDLVTFISDSTNSVCTDMGFLIDGLMPALTG 60
1 MAYRRPSLDGFIQMPKADRVKQVLAEDLVTFISDSTNSVCTDMGFLIDGLMPALTG 60
61 SHFKIAQSLAEFSLIKRIGSDFNATVATVLPVHIDRLGSDRDTVREKAQLLDLMEH 120
61 SHFKIAQSLAEFSLIKRIGSDFNATVATVLPVHIDRLGSDRDTVREKAQLLDLMEH 120
121 RVLPQALIDKLATSCFHKNAKVEEFLQITVNAHEVGTQOLSVRVYIPVCAALGDP 180
121 RVLPQALIDKLATSCFHKNAKVEEFLQITVNAHEVGTQOLSVRVYIPVCAALGDP 180
121 RVLPQALIDKLATSCFHKNAKVEEFLQITVNAHEVGTQOLSVRVYIPVCAALGDP 180
121 RVLPQALIDKLATSCFHKNAKVEEFLQITVNAHEVGTQOLSVRVYIPVCAALGDP 180
181 TVNREAAIQTLVEIKYKVGDRPLRMDVPSKAMLEOKDYOKOEBLLPSALK 240
181 TVNREAAIQTLVEIKYKVGDRPLRMDVPSKAMLEOKDYOKOEBLLPSALK 240
181 TVNREAAIQTLVEIKYKVGDRPLRMDVPSKAMLEOKDYOKOEBLLPSALK 240
181 TVNREAAIQTLVEIKYKVGDRPLRMDVPSKAMLEOKDYOKOEBLLPSALK 240
241 NTNGVGLDEADNIGLERPPTRMIRKPLHSVSSSLRPKNVNDVTDAGAVTESPES 300
241 NTNGVGLDEADNIGLERPPTRMIRKPLHSVSSSLRPKNVNDVTDAGAVTESPES 300
241 NTNGVGLDEADNIGLERPPTRMIRKPLHSVSSSLRPKNVNDVTDAGAVTESPES 300
241 NTNGVGLDEADNIGLERPPTRMIRKPLHSVSSSLRPKNVNDVTDAGAVTESPES 300
301 SFEVVPQINIFHAKMDIYKQVLTIVISDKNADMEKRVDALEKTRALLIISYHTQPFVA 360
301 SFEVVPQINIFHAKMDIYKQVLTIVISDKNADMEKRVDALEKTRALLIISYHTQPFVA 360
301 SFEVVPQINIFHAKMDIYKQVLTIVISDKNADMEKRVDALEKTRALLIISYHTQPFVA 360
301 SFEVVPQINIFHAKMDIYKQVLTIVISDKNADMEKRVDALEKTRALLIISYHTQPFVA 360
361 VOLKELSLFVDILKEELRSQVIRACCTIAVMSKTLNKLDAFCWSLLEHLINLIONSA 420
361 VOLKELSLFVDILKEELRSQVIRACCTIAVMSKTLNKLDAFCWSLLEHLINLIONSA 420
361 VOLKELSLFVDILKEELRSQVIRACCTIAVMSKTLNKLDAFCWSLLEHLINLIONSA 420
361 VOLKELSLFVDILKEELRSQVIRACCTIAVMSKTLNKLDAFCWSLLEHLINLIONSA 420
421 KVIASASTIALKYIKYTHAPKLKIYDTLNQSKDIRSTLCMLVLEFEWQTKALE 480
421 KVIASASTIALKYIKYTHAPKLKIYDTLNQSKDIRSTLCMLVLEFEWQTKALE 480
421 KVIASASTIALKYIKYTHAPKLKIYDTLNQSKDIRSTLCMLVLEFEWQTKALE 480
421 KVIASASTIALKYIKYTHAPKLKIYDTLNQSKDIRSTLCMLVLEFEWQTKALE 480
481 RNAIVLRDTLTKKSIQDADCDARRHSRYAYMAFRHHFPELADQIYGTLDIAAQRALERERE 540
481 RNAIVLRDTLTKKSIQDADCDARRHSRYAYMAFRHHFPELADQIYGTLDIAAQRALERERE 540
481 RNAIVLRDTLTKKSIQDADCDARRHSRYAYMAFRHHFPELADQIYGTLDIAAQRALERERE 540
481 RNAIVLRDTLTKKSIQDADCDARRHSRYAYMAFRHHFPELADQIYGTLDIAAQRALERERE 540
541 GGGGGGTGTGTGTAETRTYRSIRGTPTGLOKPTPSKRSISAVDTAAQAKRAQYTL 600
541 GGGGGGTGTGTGTAETRTYRSIRGTPTGLOKPTPSKRSISAVDTAAQAKRAQYTL 600
541 GGGGGGTGTGTGTAETRTYRSIRGTPTGLOKPTPSKRSISAVDTAAQAKRAQYTL 600
541 GGGGGGTGTGTGTAETRTYRSIRGTPTGLOKPTPSKRSISAVDTAAQAKRAQYTL 600
601 YSRQKPLGPNNSNOSMTGAASGLPRPLNSSGGTPATTPSVPFRGRGAVSOS 660
601 YSRQKPLGPNNSNOSMTGAASGLPRPLNSSGGTPATTPSVPFRGRGAVSOS 660
601 YSRQKPLGPNNSNOSMTGAASGLPRPLNSSGGTPATTPSVPFRGRGAVSOS 660
601 YSRQKPLGPNNSNOSMTGAASGLPRPLNSSGGTPATTPSVPFRGRGAVSOS 660
661 QPGSRSTPSTKLARDQYIGIYNGATGAI PKKASGIPRSTASSRETSPFRSGGGLMKR 720
661 QPGSRSTPSTKLARDQYIGIYNGATGAI PKKASGIPRSTASSRETSPFRSGGGLMKR 720
661 QPGSRSTPSTKLARDQYIGIYNGATGAI PKKASGIPRSTASSRETSPFRSGGGLMKR 720
661 QPGSRSTPSTKLARDQYIGIYNGATGAI PKKASGIPRSTASSRETSPFRSGGGLMKR 720
721 SMYSTGASRRTPENNNVPRPSAPARLLAQSREAHHTLGVDGDDQPPVVSQDYMRSGGMR 780
721 SMYSTGASRRTPENNNVPRPSAPARLLAQSREAHHTLGVDGDDQPPVVSQDYMRSGGMR 780
721 SMYSTGASRRTPENNNVPRPSAPARLLAQSREAHHTLGVDGDDQPPVVSQDYMRSGGMR 780
721 SMYSTGASRRTPENNNVPRPSAPARLLAQSREAHHTLGVDGDDQPPVVSQDYMRSGGMR 780

721 SMYSTGASRRTPENNNVPRPSAPARLLAQSREAHHTLGVDGDDQPPVVSQDYMRSGGMR 780
781 MGRKLMDRSDSDIDSEASSVCSEERSPDSSYTRGNKSNYSLSGSHRLDWSGTORAPDDI 840
781 MGRKLMDRSDSDIDSEASSVCSEERSPDSSYTRGNKSNYSLSGSHRLDWSGTORAPDDI 840
841 ETTIQFCASTHWSERKQGLISLTQYADGKELTQOOLKCVLDNRKPMQTHYVYSLFL 900
841 ETTIQFCASTHWSERKQGLISLTQYADGKELTQOOLKCVLDNRKPMQTHYVYSLFL 900
841 ETTIQFCASTHWSERKQGLISLTQYADGKELTQOOLKCVLDNRKPMQTHYVYSLFL 900
841 ETTIQFCASTHWSERKQGLISLTQYADGKELTQOOLKCVLDNRKPMQTHYVYSLFL 900
901 DTVELLVHANESRRGSSCLTRLPNKLGTDILNMSHKTWTLGVNHEVPTQOLK 960
901 DTVELLVHANESRRGSSCLTRLPNKLGTDILNMSHKTWTLGVNHEVPTQOLK 960
901 DTVELLVHANESRRGSSCLTRLPNKLGTDILNMSHKTWTLGVNHEVPTQOLK 960
901 DTVELLVHANESRRGSSCLTRLPNKLGTDILNMSHKTWTLGVNHEVPTQOLK 960
961 ELFRISDSTQPTTKTRIALRFLTDLANTYCKSSDPEDQOACERTYKLAQLAADQ 1020
961 ELFRISDSTQPTTKTRIALRFLTDLANTYCKSSDPEDQOACERTYKLAQLAADQ 1020
961 ELFRISDSTQPTTKTRIALRFLTDLANTYCKSSDPEDQOACERTYKLAQLAADQ 1020
961 ELFRISDSTQPTTKTRIALRFLTDLANTYCKSSDPEDQOACERTYKLAQLAADQ 1020
1021 KSMELRQARSCLVALYNLMTPOWTLILLADLPKYQDSASCIHSHRRQSQSCNSGANS 1080
1021 KSMELRQARSCLVALYNLMTPOWTLILLADLPKYQDSASCIHSHRRQSQSCNSGANS 1080
1021 KSMELRQARSCLVALYNLMTPOWTLILLADLPKYQDSASCIHSHRRQSQSCNSGANS 1080
1021 KSMELRQARSCLVALYNLMTPOWTLILLADLPKYQDSASCIHSHRRQSQSCNSGANS 1080
1081 PSSSPSSSPKPLQSPSVGPASLOSHHQLSISSTSPRSQSSVBOELLFSELDIOH 1140
1081 PSSSPSSSPKPLQSPSVGPASLOSHHQLSISSTSPRSQSSVBOELLFSELDIOH 1140
1081 PSSSPSSSPKPLQSPSVGPASLOSHHQLSISSTSPRSQSSVBOELLFSELDIOH 1140
1081 PSSSPSSSPKPLQSPSVGPASLOSHHQLSISSTSPRSQSSVBOELLFSELDIOH 1140
1141 NIQTSSEIRHCFQGYOTALAPNGFNGLQYHDQGOQDSCASLSNSKTQSSANTQSN 1200
1141 NIQTSSEIRHCFQGYOTALAPNGFNGLQYHDQGOQDSCASLSNSKTQSSANTQSN 1200
1141 NIQTSSEIRHCFQGYOTALAPNGFNGLQYHDQGOQDSCASLSNSKTQSSANTQSN 1200
1141 NIQTSSEIRHCFQGYOTALAPNGFNGLQYHDQGOQDSCASLSNSKTQSSANTQSN 1200
1201 TPESATWRLNLEBERTTQNAKSPTDAKITYYSINAAENGELILANLMESEVVAALT 1260
1201 TPESATWRLNLEBERTTQNAKSPTDAKITYYSINAAENGELILANLMESEVVAALT 1260
1201 TPESATWRLNLEBERTTQNAKSPTDAKITYYSINAAENGELILANLMESEVVAALT 1260
1201 TPESATWRLNLEBERTTQNAKSPTDAKITYYSINAAENGELILANLMESEVVAALT 1260
1261 LTKQOPELLOSLTNLNGICIKGNCCLPKHFRSIRMLNLTLEAHTVIVAGLHVL 1320
1261 LTKQOPELLOSLTNLNGICIKGNCCLPKHFRSIRMLNLTLEAHTVIVAGLHVL 1320
1261 LTKQOPELLOSLTNLNGICIKGNCCLPKHFRSIRMLNLTLEAHTVIVAGLHVL 1320
1261 LTKQOPELLOSLTNLNGICIKGNCCLPKHFRSIRMLNLTLEAHTVIVAGLHVL 1320
1321 KIMSNNRNNMHLFELLILKIIQCYOHSKEALRDISMI PRIAPSLPLDSINIVPV 1380
1321 KIMSNNRNNMHLFELLILKIIQCYOHSKEALRDISMI PRIAPSLPLDSINIVPV 1380
1321 KIMSNNRNNMHLFELLILKIIQCYOHSKEALRDISMI PRIAPSLPLDSINIVPV 1380
1321 KIMSNNRNNMHLFELLILKIIQCYOHSKEALRDISMI PRIAPSLPLDSINIVPV 1380
1381 IATGEPTNCAIKILLETVEHNGSETDAHLDIVPNNLARSADDTQSMVRKAAPFCIVK 1440
1381 IATGEPTNCAIKILLETVEHNGSETDAHLDIVPNNLARSADDTQSMVRKAAPFCIVK 1440
1381 IATGEPTNCAIKILLETVEHNGSETDAHLDIVPNNLARSADDTQSMVRKAAPFCIVK 1440
1381 IATGEPTNCAIKILLETVEHNGSETDAHLDIVPNNLARSADDTQSMVRKAAPFCIVK 1440
1441 LYFVLGEKVKPKLSVLNPSKVRLLNVIIEKORNCISGGSSSTKNSAASS 1492
1441 LYFVLGEKVKPKLSVLNPSKVRLLNVIIEKORNCISGGSSSTKNSAASS 1492
1441 LYFVLGEKVKPKLSVLNPSKVRLLNVIIEKORNCISGGSSSTKNSAASS 1492
1441 LYFVLGEKVKPKLSVLNPSKVRLLNVIIEKORNCISGGSSSTKNSAASS 1492

RESULT 2

AAAB29634 standard; protein; 1526 AA.

AAAB29634;

21-FEB-2001 (first entry)

Human polliosis-associated gene 373-encoded protein, SRQ ID NO:2.

Human; polliosis-associated gene 373; IgE; immunoglobulin E;

cedar pollen allergy; T-cell; reduced expression; detection; diagnosis;

drug screening; allergic disease.

Homio sapiens.

W0200065046-A1.

02-NOV-2000.

26-APR-2000; 2000MO-JR002730.

PR 27-APR-1999; 99GP-00120489.
 (GENO-) GENOX RES INC.
 XX Nagasu T, Sugita Y, Kaishwabara T, Oshida T, Obayashi M, Gunji S;
 PI Obayashi I, Imai Y, Yoshida N, Ogawa K, Matsui K;
 XX WPI: 2000-687339/67.
 DR N-PSDB; AAC64201.
 XX
 PT Pollinosis-associated gene 373 undergoing significantly low expression in
 PT subjects with high cedar pollen-specific immunoglobulin-E levels, useful
 PT in diagnosis of allergic diseases and screening drug candidates.
 PS Claim 1; Page 58-69; 80pp; Japanese.
 CC The invention relates to the human pollinosis-associated gene 373 which
 CC exhibits significantly reduced expression in the T-cells of individuals
 CC with high cedar pollen-specific IGE (immunoglobulin E) levels. The gene
 CC was isolated from T-cells from individuals allergic to cedar pollen using
 CC the differential display method. The invention also relates also relates
 CC to the protein encoded by pollinosis gene 373; expression constructs and
 CC host cells comprising pollinosis-associated gene 373 nucleic acids;
 CC pollinosis-associated gene 373 primers and probes; antibodies against the
 CC protein encoded by the gene; methods of detection of pollinosis-
 CC associated gene 373 nucleic acids; and a method of diagnosis of allergic
 CC diseases via the detection of pollinosis-associated gene 373 nucleic
 CC acids. The invention additionally encompasses methods of screening drug
 CC candidates for the treatment of allergic disease by measuring the
 CC expression of pollinosis-associated gene 373 in pollen antigen-stimulated
 CC T-cells in the presence of a test compound relative to a control.
 CC Pollinosis-associated gene 373 is useful in the diagnosis of allergic
 CC diseases and in the screening of drug candidates for the treatment of
 CC such diseases. The present sequence represents a protein encoded by human
 CC pollinosis-associated gene 373
 CC
 XX
 SQ Sequence 1526 AA:
 Query Match 22.4%; Score 1707; DB 3; Length 1526;
 Best Local Similarity 30.1%; Pred. No. 6,3e-119; Indels 206; Gaps 45;
 Matches 485; Conservative 300; Mismatches 618;
 QY 5 KPSSDGLIQMPKDMRVQVLAEDLVFL-SDDTNSIVCTDGMFL--IDGLMPWLTG 60
 DB 1 EPRSMHYCAQYQKQDVGRLVGGELLYLGAPALISDLEEDLRLGKTVADALGWGVS 60
 QY 61 SHFKIAQSLKFAFSELIKRLGSDFNAYTATVLPVHVIDRIGSDRYREKQALLDLMEH 120
 DB 61 SNYKVSIMQLEILSAFVDRILSTRFSYVAMVIVALIDRMGDADKDKRDEAQTLLIKMD- 119
 QY 121 RVLPPQALIDKLATSCFKRKAQVREPIQTIYNLHEGYTQOLSRYVYIPVICALGDP 180
 DB 120 QVAPFMYIWEOLA-SGFQKHFRSRGVCCLLETINIFGAQPLVYSKILPHLICILFGDS 178
 QY 181 TYNVREAAQTQVEIKVGDRLRPLRMDVPAKSLMLDQKEDQVQEGULLPSALK 240
 DB 179 NSQVRDAALAIYVIRAHGKRMOLYKR-GIPPARLEWIFAKFDEVQSSGGMILSVCK 237
 QY 241 NTNGNGVGLDEADNIGLRERPT-----RMIKRPLHSAVSSSLRPKPVNDVT-- 287
 DB 238 DKS-----FDDDESVD-GNRPSAASAFKVPAPKTSGNPANSARKPGSAGSPVGASKE 291
 QY 288 GDGAGVATMESFESSFEVYPOALNFHAKMDMDIKOVUVIISDKNADMEKRVDAKKIRAL 347
 DB 292 GGAGAVDEDDFIFAFDVPISIQYSSRELETANKIRELISDDKMDQAMAKKIRSL 351
 QY 348 LI---LSYHTQPOFVAQVQLKELISFVDILKEELRSQVIREACTITIAVSKTLKTLADP 404
 DB 352 LVGAAQYDQCFQRLRLDLDAKLS-----AKDLRSQVREACTIVASLSTVLGNKFDHG 406
 QY 405 CWSLIEHILNLIQNSAKVIASTIALKYYIKYTHAPKLIKITYDTLNOGSKMDIRSLC 464
 DB 407 AEALVPLTFLNLPVPSAKVMAISGCAIRFIIRHTRVPLIPLITSNCT-SKSVFVRRRSF 465

QY 465 ELMTVLLFEEMOTKALERNATVLRDLTKSISGDADCDARRHSRYVMAFRHPPELADQY 524
 DB 466 EFLDLLLOEMQTHSLERNAVALVETIKKGIHDADAEARVEARKTYGDRNHPGEAETLY 525
 QY 525 GTLIDIAQPALEREREGGG-----GGTGTGTGTAPEPTR 559
 DB 526 NSLEPSYQKSLQTLTKSSGSVAALPQSDRSSSSQESLNRPPSSKWTANPESTVAGRVSA 585
 QY 560 TVSRIGRPTGTLQKTPSWRSISAVDTAAQRAKYRAQVTLVSRQKPLGPNNSNOAS-- 617
 DB 586 GSKKASLPQSLQ-----KRSNDIDVNAAGAKAHHAAGSVRSGR-LGAGALNAGSYA 638
 QY 618 -----MTGAAAGSLPRPLNSNGTPTATPGSVTPPRRGAG---VQSQPGSS 666
 DB 639 SLEDTSDKLDGTASBGRVRAKLSA-----PLAGMNAKADRGSRRTMVSQSGPSHS 693
 QY 667 TSP-----STKLRDQYIGNYRGATGALPKKASGIPSTIASRSTSPTR---SGGGLM 718
 DB 694 GSGRVLTTTALSTVSSGVQVLYVNSASA--OKSKIPRSQCSREASPSRLSVASRSI 751
 QY 719 KSMSTGA-----GSRTPERNNVPRPSAP-----A 745
 DB 752 PRPSVQSGSREASRESRSDTSVNSFQPLASHHSRTGALYAPEVYAGSGGYGISQS 811
 QY 746 RLTAQREAEHTLVGDDQPDVYSGDYRSGGMRGKLMGEDESDDIDSEASSVCSER 805
 DB 812 SRLSSVSAMRVLNTSD--VEEAVADALKKPARRR-YESYGMHSDDANSDASSACSER 868
 QY 806 SPDSYTRGNKSNVSLSGSHTRLDWSTORAPPDITIIIOFCASTHMSERKGLISLTQY 865
 DB 869 SYSS-----RNGSI-----PYMKQTDAVEALNRKASNNSEKRGELGIONL 912
 QY 866 LAQKELTQOQKCVLDMPEKMFMDPTK--VSLEFLDTVTELLVHANETSRNGSSCLT 924
 DB 913 LKQRTLSRVELKRLCETFRMRADPHGKVPFMSLETLTVDFIQVHDDL-QDWLFVILT 971
 QY 925 RLFWKLTGDTLNSHMKIWKTLQVHNEYPTQQLKELPRIISDSTOTPTTKRIALIRF 984
 DB 972 QLKMKADLLGSQAVQKALQADVTRESFPNDQFNIMLFMTVDOTQTPSLKVALILKY 1031
 QY 985 LTLDLANTYCKSDPSD--OSQACERYVLAQLAALQKMEARSQASRCVALYNLNTP 1042
 DB 1032 IETLA-----KOMD-PGQFINSSBETRLAVSRVITWTTPKSDVKAQSVLISPELNTP 1086
 QY 1043 QMTLLADLPVYQDASRCSHSMRROSQSCNGANSPSSPLSSSPK-----PLQ 1095
 DB 1087 EFTMLGLPPTFDGATKLLHNHR---NTNGTQSSMGSPLRPTRRPAMWSPLT 1142
 QY 1096 SPASGPFASLQSHHQLSISSTSPRSQSYE-----QELLFSELDIQHNIOKTSEE- 1148
 DB 1143 SPYNTSONTLPSPAFVDYTEMNSEDYSSLRGTALIONFPSQEDMNEPLKDSKDD 1202
 QY 1149 --IRHCFGQYQYOTALPENGHGLQYHDQGDQSCASISNSKQSSANTQSN---TP 1202
 DB 1203 DGDMSCGG---PGMSDPRAGDADTDSQOTALDNKASLSHSPHTSSPBRDYNPNYNSD 1258
 QY 1203 ESATKRLNLERERTQOAKSPTDQAKVITYSINNAENGEBLLASNLMESEVVRVALTLT 1262
 DB 1259 SISPNKSLKXAMFDDADQFPD-----LSLHSD-----LVAELEKLS-----N 1301
 QY 1263 KDQPELLOTSITNLGICIKGNCELPNKHSINRMILNLTLEAHTDVVINGLHLSKI 1322
 DB 1302 HNERVEERKIALYELMKLQESFVWMBHFITILLLETGLGDEPPTRALAKVLRBI 1361
 QY 1323 MRSNKRHNMMHLELILIKITIQYQH-SKALRQIDSMIPRIAPSLPDLGINVNPYI 1381
 DB 1362 LKHQPAR--FKVYAEITVWKTLEAKDPRKEVVAEAAEASVLAISISPEQICVLCPII 1419
 QY 1382 ATGEPTNICALKILLEVEHSGSETDAHLDIVFPNLARSADTQSMVKAAVFCIVKL 1441
 DB 1420 QTADEVINLAIMQTVIERVSKETLNLILBEIMPGLQGYDNSSSVRKACVFCIVAV 1479

QY	1442	YFVLGEENKPKLTVNPKRLLNLYIEKORNCISGGSTGNSSAAS	1490
DB	1480	HAVIGDE-LKPHLSQLTGSKMKLLNIYI-KRAQTSGGADPTTIVSGQS	1526
		RESULT 3	
ID	ADF60378		
AC	ADF60378	standard; protein; 1298 AA.	
DT	12-FEB-2004	(first entry)	
DE	Human config polypeptide sequence SEQ ID NO:2745.		
XX	biological activity; genetic engineering; hybridisation probe; oligomer;		
XX	primer; chromosome mapping; gene mapping; recombinant protein production;		
XX	human.		
OS	Homo sapiens.		
XX	MO2003080795-A2.		
PN	02-OCT-2003.		
XX	09-AUG-2002; 2002MO-US025465.		
PF	09-AUG-2001; 2001US-0311261P.		
FR	(HYSE-) HYSEQ INC.		
PA	Tang YT, Yang Y, Wang Z, Weng G, Ma Y;		
XX	WPI: 2003-876918/81.		
DR	N-PSDB; ADF59926.		
XX	New polynucleotides, useful as hybridization probes, oligomers or		
PT	primers, for chromosome or gene mapping, for the recombinant production		
PT	of proteins, and for generating antisense DNA or RNA.		
XX	Example 3; SEQ ID NO 2745; 571bp; English.		
PS	The present invention describes isolated polynucleotide sequences (I),		
CC	which encode polypeptides (II) with biological activity. Also described:		
CC	(1) a vector comprising (I); (2) an expression vector comprising (I); (3)		
CC	a host cell genetically engineered to comprise (I) which is operatively		
CC	associated with a regulatory sequence that modulates expression of (I) in		
CC	the host cell; (4) a polypeptide (II) encoded by (I); (5) a composition		
CC	comprising the polypeptide of (4) and a carrier; (6) an antibody directed		
CC	against the polypeptide of (4); (7) detecting (I) or the polypeptide of		
CC	(4) in a sample; (8) identifying a compound that binds to the polypeptide		
CC	of (4); (9) producing the polypeptide of (4); and (10) a collection of		
CC	polynucleotides comprising at least one of the polynucleotide sequences		
CC	(I). The polynucleotides (I) can be used as hybridisation probes,		
CC	oligonucleotides or primers, for chromosome or gene mapping, for the recombinant		
CC	production of proteins, and for generating antisense DNA or RNA. The		
CC	present sequence represents a human config polypeptide sequence, which is		
CC	used in an example from the present invention.		
XX	Sequence 1298 AA:		
QY	Query Match	17.8%; Score 1352.5; DB 7; Length 1298;	
	Best Local Similarity	29.9%; Pred. No. 3.1e-92;	
	Matches	415; Conservative 235; Mismatches 496; Indels 24; Gaps 47	
QY	236 PSALKNTNGNGVGLDEADNIGLRERPTMKRPLHSAVSSSLRPKNVNDVTGAGAVTM	295	
DB	15 PSSERN-----VGGTTRRLG-----SSTLGSKSA--AKEGAGAVDE	50	
QY	296 ESFESSFEVPOINIFHAKMDMDIYKQVLVIISDKNADMEKRVDAKTKIRALLIISYHTQ	355	
DB	51 EDFIKAFPDVAVVOYISRDDEESINKRKREILISDQKHMEQVNAKTKIRSLLAGAAEY	110	

QY	356	PQFAVAVOLKELSLASFDVILKREILRSQVIREACTITAMSSTLRKRLAFCMSILIEHINTL	415
Dd	111	DNFQO-HURLIDGAF-KLSADOLDSQVVRREACTILGHLSSVLGKPFDHGAEMIPPTIFNL	168
QY	416	IONAKVYIASASTALKYIIRKYTAPAPILKITYDTLNLQSKSKODRSTLCEIMLVLFEEWQ	475
Dd	169	IPNKAQIMATYSQVAVAVLLIRHHTIIRPLIIVITNSNC-SKAVAVRRRCOFELDDLLOQEWQ	227
QY	476	TKALERNATVLRDTLTKKSGDADCDARHSRIYAWAFRRHFPPELADQIYGTLDIAORAL	535
Dd	228	THSLERHISVLAETIKKGIHDADSEARIEARKCWGSHFSFEAEHLVHTLEBSYOKAL	287
QY	536	EREHEGGGGGTGCTGCTAPETRIITYRIRIGTETLOKPTPMSRISIAVDPAQAQAKVR	595
Dd	288	QSHLK-----NSDSIVSLPOSDDSSS--SSQESLRLPLSARKSPTGTTSRISTVSTK	338
QY	596	AQYTLVSRQRPFLGPNNSNOASMTGAAGSLPRPLNSNGCTP-----ATTAPGV---	647
Dd	339	SVSTTSSLQR-----SRSDIDVNAALAA--KSKVSSSGTTTFSSAALPPGSAVSL	388
QY	648	-----TPRBRGAG--VSQSQ-----PGSRSTSP	669
Dd	389	DGTTTKAEGRIIRTRQSGSGATVWASTPDRNGRSRAKVVSQORSRSGADAGSFRSSP	448
QY	670	STKLARDYGI-QNYTNG-ATGAIPIKASGIPRSTASRSTSPRSG---CGLMKREMY	724
Dd	449	GKLGSYGIGLGTGSSSGRPVTPSPSEKRSKIPRSGCSRETSPRIGLARSRIPRSMWS	508
QY	725	TGAG-----SRTPPERNNVVRPSAPARLLAQSREABHTLGVDDQO-PDYVSGDVNRSGG-	778
Dd	509	QCGSRDTSRSGSRDTSAPRGFPL-----DRGLQOPGRIPQSVAMRVLSTSD	538
QY	779	-----MRGKRLMGRDESDDIDSEASSVCSERSFDSYTRGNKSNYSLSGSHT	826
Dd	559	LEAAVADALKKPVRRREPYGWTSDDDANSDAVSCEERSYGS--RNGGIPHY-----	609
QY	827	RLDWSTORAPFDIETIIOFCASTHNSERKDGILSTQYVADGKELTQOOLKCYLDMFRK	886
Dd	610	-----LRQTEBVAAEVLNHCASSNMSERKEGLIGLONLKSORTLSRVELKICEFTR	662
QY	887	MEMDTHTK-VYSLFLDVTVELIVHANETSRNSSSCITLTFPKLG-TDLNMSHSMKIM	944
Dd	663	MFADPRSKRVFSMFELEIVDFIIHKODL-QDMLFVLLTOLKNGEADLLGSAQAVOK	721
QY	945	TLOVNEHFPYTOQLKELFRJISDSIQTPPTKTRIALIREFLTDLANTYCKSSDPSPD-Q	1002
Dd	722	ALDVTDRSPFDQOENILMRFIYDQIQTPMLKVAAILKITYESLA---RQMD-PTDFVN	776
QY	1003	SQACERVLKLAOLADQKSMELRSQARSGVLVNLNTPQMTLLADLPKVYODASRSC	1062
Dd	777	SSETRILAVSRIIIMTTEPKSSDVKRAQAQIVLISLFEINLTPBEFTMLGALPKTQDQATKL	836
QY	1063	IHSHMRQSGSCNSGANSPPS-----SPLSSSPKPLQSPS-----VGPPA-----	1103
Dd	837	LHNHLKNS--NTSVGSPBNTTIGRTPSRHTSRTPLISPTNCSHGGLSPSLKMGAD	893
QY	1104	SLOSHHOLISSTSPRSQSSVEOELLSPSELIDIO-HNIQKTSEERIRHCFGQOYOTALA	1162
Dd	894	GLAKHPRPFGQPSIPTAPBHKALRSYSBMDYTEMN--NSEELYSLREVTE-AIE	950
QY	1163	PNGFNHLYQHDGQOD--SCASSNSKTOSSANTTQSGNTP-BSATMRLDNLERERT	1218
Dd	951	KSFPSRQEDLNEBFIKRDGKKCECIVSRDGAASPATEGRGGSSEVGGRTALDNKTSILNT	1010
QY	1219	QNKS-----PTDDAVITIVSINMAENGL-----IL	1245
Dd	1011	QPPRAAPRGPRARDYNFYSDA-----INTYDKTALKEAVPDDMEQLRDVPIHSDLV	1064
QY	1246	ASNIMSESVVRVALTLTKDQVELLO-TSLTNLIGICIKGNCBLPMKHGRSIRMRMLNIL	1304
Dd	1065	ADLLKELSNHNEVEERKQALLLELKITREDSIGVA-----BEHKTILLLELTL	1115
QY	1305	EAEHTDVVIAGLHVLSKIMSRKMRHNMHFLLELLIKIIOQYOS-KEALRDISMTIPR	1363

```

Db      1116 GDKHSHRALALFVLEKREILLENOPAR--FKNYAELTITMKTLEAKHDSHKEVVPRAEPAAT 1173
Qy      1364 IASLPLDLSTINIVNPIATGEPTNLCATKILETEHHGSEITDAHLIDVPPNLAESA 1423
Db      1174 LASHIPEOCIKVLCPIOTADPFINLAIMOTKVERAKSLLOLVDIIPGLLOG 1233
Qy      1424 DPTQSWRKAAFCIVLTVLGESEKVPKXLSVLPSPKXVLLNVYIEKORNCISGGGSSST 1483
Db      1234 DNESSVRKASVFCVLAIVSIGED-LKPHLAQLTQSKMKLNLNLYIKR-----AQT 1283
Qy      1484 KNSSAASSS 1492
Db      1284 TNSNSSSS 1292

RESULT 4
ADB79873
ID ADB79873 standard; protein; 1324 AA.
XX
XX ADB79873;
XX
XX 04-DEC-2003 (first entry)
XX
XX Rat putative KIAA0627, SEQ ID 113.
XX
XX Analgesic; pain; streptozocin-induced diabetes; rat.
XX
XX Rattus norvegicus.
XX
XX EPI279744-A2.
XX
XX 29-JAN-2003.
XX
XX 26-JUL-2002; 2002EP-00255249.
XX
XX 27-JUL-2001; 2001GB-00018354.
XX
XX 07-FEB-2002; 2002GB-00002910.
XX
XX (WARN ) WARNER LAMBERT CO.
XX
XX Brookbank RA, Dixon AK, Lee K, Pimock RD;
XX
XX MPI: 2003-395407/38.
XX
XX N-PSDB; ADB79874.
XX
XX Use of isolated gene sequences and encoded polypeptides that are
XX upregulated in the spinal cord in response to streptozocin-induced
XX diabetes for screening compounds for the treatment of pain, or for
XX diagnosing pain.
XX
XX Claim 1; Page 204-208; 334pp; English.
XX
XX The present invention relates to nucleotide sequences which are useful in
XX the screening of compounds for the treatment of pain, or for the
XX diagnosis of pain. The nucleotide sequences are up-regulated in the
XX spinal cord in response to streptozocin-induced diabetes. The present
XX sequence was used to illustrate the invention.
XX
XX Sequence 1324 AA:
XX
XX Query Match 17.7%; Score 1348; DB 7; Length 1324;
XX Best Local Similarity 29.6%; Pred. No. 6.9e-92;
XX Matches 404; Conservative 233; Mismatches 523; Indels 204; Gaps 39;
Qy      237 SALK-----NTNGNGVGLDEADNIGLERPTRMIRKPLHSVSSSLRPKNVNDVTDGAGA 292
Db      55 SAFKVPAPKTSGN-----PANSARKP-----GSAGGPKVGASKEGAGAGA 94
Qy      293 VTMESESSFEVVPOLNTHAKMDMDIYQVLIISDKNADWEKRYDALKTIKRIALLI--- 349
Db      95 VVEDDFIKAFITDVPFIOIYSSRELLETINKIREIISDDKHWDQDANALKTIKRIALLVAGA 154

```

```

Qy      350 LSYHTQPFVAVOLKELSLSFVDLIKELRSGVIREACTITTAVMSKTLRNKLDACWML 409
Db      155 AAYDCEPQHLRLDGLAKLS-----AKDLRSGVIREACTITVAHLETLVGNKEDHGAEATV 209
Qy      410 EHLNLIIONSAKVIAASTALIKYITIKTHAPKLKITDTLINOKSKDIRETLCLEMLV 469
Db      210 PTLFNLVPASAKVMAATSGCAIRFIIRHIVRLPLITISNCT-SKSVVRRRSFEFLDI 268
Qy      470 LREEMOTKALENNATVLTDLTKKSIQDADCDARRSRVYVMAFRHFPPELADQIYGLDI 529
Db      269 LIQEWQTHSLERHAALVETLIKKGIHDADEARVEARKTYMGLRNHFPPEAEITVLSLP 328
Qy      530 AAORALEREREGGG-----GGTGCTGTATPRTTRVSI 564
Db      329 STOKSLQITLKSSGSVALSPQSDRSSSSQESLNRPFSKMSVTANSTVAGVSAAGSVXA 388
Qy      565 GRTPGTLQKPTSMRSISAVDTAAARAKVRAQYITLSPQRKPLGNPNNGQS----- 617
Db      389 SSLPGSLQ-----NSRSDIDVMAAAGAKAH-HAAGQSVRRGLGAGALNAGVSLBDT 441
Qy      618 ---MTGAAGSGSLPRRLNSNGGFPATTPGAVTPRPRGAG---VSOGQPGSRSTSP-- 669
Db      442 SDKLDGTASEDEGRVRAKLSA-----PLAGMNAKADSRGRSRTKMYSQSGPGSRSGSPGR 496
Qy      670 ---STKLDPQYGGIGNYVRGATGALPKKASGIPRTASRETSPTP---SGGLMKRSMY 723
Db      497 VLTTLTSLVSSGVQVLYNSASA--QKRSKIPROGCSREAPPSRLSVARSRIIPPSV 554
Qy      724 STGA---GSRRTPERNNPVPRSAV-----ARLQAQ 750
Db      555 SQGCSREAGSSSRDTSVPYSPQPLASRHSRSTGALYAEVYGAGPGYGISQSSRLSS 614
Qy      751 SREAEHTLVGDDGQGDYVSGDYMRSGRMGRKLGKRDSDIDSEASVCSERSPDS 810
Db      615 SVSAMRVLTGSD--VEEAVADALKKPARR-YESIGMHSDDANSASACERSYSS- 670
Qy      811 YTRGKSNYSLSGSHTRLDWSTORAFDDIETIIQFCASHTWSEKDKGLISLTQYLAQK 870
Db      671 -----RNGSI-----PTMROTEDVAEVLNRCASSNMSRKEGLLGLQNLNQR 715
Qy      871 ELTQOOLKCYLDMFRKXMDTHTK-VYSLFLDVTTELIVHANETSNGSSCLTRLENK 929
Db      716 TLSRVELKRLCEIFTRFADPHGRKVFSELETLVDFIQVHKDL-ODWLFTLLTOLKK 774
Qy      930 LGTDLNSMSKTKMLQVYHVEYFPQLOKELPRIISDSQPTTTRIALIRFLDIA 989
Db      775 MGADLLGSVQAKVQKALDVRESFPNDLFNILRFTVDTOTPISLKVAAILIKIETLA 834
Qy      990 NTYKSSDPFSD--SQACERTVLKLAQLAADOKSMELRQARSCLVALYNLTPQMTLL 1047
Db      835 ---KQMD-PEDFINSEETRLAVSRVITWTTERKSSVRAKAGVLLISFELNTPPEFTML 889
Qy      1048 LADLPKYODSARSCISHNRROSQSCNSGANSPPSSSPK-----PIQSPSVG 1100
Db      890 LGLAPKTFQOGATGLLNHLR-----NTNGNGQSMGSPLTPTPRSPAMWSPPLTSPNT 945
Qy      1101 PPSALOSHNLQSLSSISPSRSOSV-----OELLFSSRLDIOHNOKTSEE---IRH 1151
Db      946 SQNTLSPSAPDYDTENNNSEDIYSLRGVTEAIONFSRSDENMEPLKRSKODDGSIM 1005
Qy      1152 CFGGOYOTALAPNGFNGLQYHDOQODSCASLSNSKTQSANTTQSN---TPESATM 1207
Db      1006 CGG---PGMSDPPRAGDATTSSQTALDNKASLSLHSMPTSSPSRDNPNVYSPSISPF 1061
Qy      1208 RLDMLEERTQONKASPTDAAKVITVSINNAENGELLIASLMSSEVVRVALTLTKQDPV 1267
Db      1062 NKSLAKEMAFDDADQFPDD-----LSLDHSD-----LVAEELKELS-----NHNRV 1104
Qy      1268 ELDTSLNLGICIKGAGCELPNKHFRSIMMNLNIEAEHTDVYIAGIHLYSKIMRSNK 1327
Db      1105 EERKIALYEMLKLTQOESFYSWDEHFKTLLLETJGDKPPTRALAKVLKREILRQIP 1164
Qy      1328 MRHMMHFLLEILIKTIQCYOH-SKEALRIDISMIPRIAPSLPLDLSINIVNPIATGEF 1386

```

```

Db      1165 AR--FKYVAELTVNKTLEAHDKPHKEVVRSAEASVLAISISEQCICVLCPIIQADY 1222
Qy      1387 PYNCAIKILEVEVHGHSEITDAHLDIVFNLARSADDDTOSMWRKXAVFCIVLYVLG 1446
Db      1223 PINLAIKMQKVIERYSKETLNLPEINBGLIQGIDNSESSTRKACVFLVAVHAYIG 1282
Qy      1447 EEKYKPKLSVLNPSKRVLLNVIYIEKORNCISGGSSSTKNSAAS 1490
Db      1283 DE-LKPHLSQLTSGSKMKLNLNYI-KRAQTGSGGADPTTVDVSGQS 1324

RESULT 5
ADK07258
ID ADK07258 standard; protein; 1324 AA.
XX
AC ADK07258;
XX
DT 21-APR-2005 (first entry)
XX
DE Cyclin-dependent kinase modulation biomarker SEQ ID NO 1823.
XX
KM cytoeatic; cyclin-dependent kinase; cdk; biomarker.
XX
OS Homo sapiens.
XX
PN MO2005012875-A2.
XX
PD 10-FEB-2005.
XX
PF 29-JUL-2004; 2004MO-US024424.
XX
PR 29-JUL-2003; 2003US-0490890P.
XX
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;
XX
DR WPI; 2005-163068/17.
XX
N-PSDB; ADK07257.
XX
PT Biomarkers useful for predicting or determining the response of a mammal
XX
to a cancer treatment comprising administration of a modulator of cyclin-
XX
dependent kinase activity.
XX
PS Claim 5; SEQ ID NO 1823; 141pp; English.
XX
XX
This invention describes a novel method of predicting or determining
whether a mammal will respond or is responding to an anti-cancer agent
that modulates cyclin-dependent kinase (cdk) activity. The method
comprises measuring the level of one or more biomarkers selected from
2774 biomarkers given in the specification (nucleotide sequence SEQ ID
NO:1246 (Genbank EST W28729) is especially preferred). The method of the
invention is utilized in a kit for determining or predicting whether
a patient would be susceptible or resistant to treatment by an agent
modulating cdk activity. The invention also describes a method for
utilizing individualized genetic profiles for treating diseases and
disorders based on patient's response and molecular level, specialized
microarrays comprising the biomarkers described, antibodies directed
against the biomarkers and a cell culture model to identify biomarkers.
The cdk modulator is preferably N-5-[(5-(1,1-dimethyl-2-oxazolyl)methylthio)-2-thiazolyl]-4-piperidine carboxamide, 0.5-L-
tartaric acid salt. Note: The sequence data for this patent did not form
part of the printed specification, but was obtained in electronic format
directly from WIPO at ftp.wipo.int/pub/published_pat_sequences. This
sequence represents a biomarker used in the method of the invention.
XX
SQ Sequence 1324 AA;

```

```

Query Match 17.7%; Score 1348; DB 9; Length 1324;
Best Local Similarity 29.6%; Pred. No. 6.9e-97;
Matches 404; Conservative 233; Mismatches 523; Indels 204; Gaps 39;

```

```

Qy      237 SALK---NITNGVGVLDEADNIGLEREPTRMIKRPLHSAVSSSLPKPVNDVTDAGA 292
Db      55 SAKFVAPAKTSGN-----PANSARKP-----GAGGPKVCGAGKEGAGA 94
Qy      293 YIMSEFSEFVPPVPOINIFHAKMDDIYKQVLIISDNKADMKRDALAKTIALI--- 349
Db      95 VDEDDFTKAFPTDPSIOIYSRRELETLNKEILISDDHDWQORANALKTRSLVAGA 154
Qy      350 LSYHTQPOFAVAVOKELSLSFVDILKEELRSQVIREACTIIAYMSKTLBNKLDAFCMSIL 409
Db      155 AQYDCFPQHLRLLDGALKXS-----AKDLRSQVREACTIVAHLSVLGNKPDHGEAIV 209
Qy      410 EHLINIQNSAKYIASASTIALKYIIKYTHAPKLYITDTLNQSKDRISTLCMLVL 469
Db      210 PTLFNLVPSAKMAYATSGCAIRFIRHTHVPLRPLITSNCT-SKSVPRRRSFEFLDL 268
Qy      470 LPEENQTKALERNATYLRDTLKKSIGDADCDARHRHRYMYAFRRHFPPELADIYGTLDI 529
Db      269 LQEMQTHLERHAVALVETIKKGIDHDAEAKVEARKTYMGLRNHFPGBAETLYNSLEP 328
Qy      530 AAOBALEREREGCG-----GGTGTGTGAPETRRRTVSRI 564
Db      329 SYOKSLQTYLKSGSVASLPQSDRSSSSQESLNRPFSKMTANSTYAGRVASAGSKA 368
Qy      565 GRTPGTLQKPTPSMRSISAVDTAAAGRAVRAQYTLYSRQKPLGNNSNQAS----- 617
Db      389 SSLPGSLQ-----RERSDIDVNAAGAKAH--HAAGQSVRRGRGLGALNAGVSLLED 441
Qy      618 ---MTGAAGSGSLPRRLNLSNGGTPATTPGSVTPPRRAG---VSOGQPGSRSTSP-- 669
Db      442 SDKLDGTASDEDRVRKLSA-----PLAGMGAKADSRSRSTKRVYSSQSPGSRSSPCR 496
Qy      670 ---STKLRDQYGAIGNYRGATGATGAIKPKASGIPRSTASRETSPTR---SGGLMRSMY 723
Db      497 VLTITLSTVSSGVQVRLVNSASA--QKRSKIPRSGCGRERASPSLSVARSRIRIRPEV 554
Qy      724 STGA---GSRTPPERNNPVPSAP-----APLLAQ 750
Db      555 SGCSREASRESRSRDTSPVRSFQPLASRHSRSTGALVAPEVYASGPGYGISQSRLSS 614
Qy      751 SREAEHTLGVGDGQDGYSGDYMSSGMRMKLMGRBESDIDSEASVCSERGFDS 810
Db      615 SYANRVLNTGSD--VEEAVALKKPARR--YESYQMSDDANSDASACERSYS- 670
Qy      811 YTRGNKSNYLSGSHRTLDWSTORAFPDDETIIIOFCASTHMSERKDGHSITQVYADGK 870
Db      671 -----RNGSI-----PTYMQTEDEVAVLARCASSNMSERKEGLGLGNLKNQR 715
Qy      871 ELTQOQKCVLDMFRKMPMDYTK-VSLPDTVTYELLIVHANEYSRNGSSCLTRLPNK 929
Db      716 TLSRVELKELCEIFTRMFADPHGKRVSMFLETLVDFOVHKDDL-QDWLFVLLTOLKK 774
Qy      930 LGTDILNHSKIKWTLQVNHVEFPQQLKELFRISSTOPTTKTIALIRFLTDA 969
Db      775 MGADLLGSQAVQVQKALVDTRSPFNDLOFNIMRVTYQOTPSLKVAVALIKYETLA 834
Qy      990 NTYCKSDPSPD--OSQACERTVLKLAQADQKSELSQARSCLVALYNLTPTOMTLL 1047
Db      835 ---KQMD-PGDFINSSERLAVSRVITWTTERKSPDVAKAQSVALISLPELNTPEFTWL 889
Qy      1048 LADLPKVYDASRCSIHSHMRQOSCNCSGANSPPSSSPK-----PLGSPSYG 1100
Db      890 LGLLPFTPDGATKLLHNHLR---NTNGGTOSMSGSLTRPTPRSPAMWSLTPTYT 945
Qy      1101 PRASLOSHHQSISTSPRSQSSVE-----QELFSELDIQNIQKTSSE---IH 1151
Db      946 SONTLSPSAFYDVTENWSEDIYSSLRGVTAIQLNPFSSQDEWNEPLKRDSDKDDGDSM 1005
Qy      1152 CFGGQYOTALADPNFGHLYQYHDQGOQSCALSSNSKTOSANTTOSN---TPESATM 1207
Db      1006 CGG-----PMSDPRAGGADATDSSQALDNKXSLHSMPTHSSPRRDVNPVYSDSISPF 1061
Qy      1208 RLDNLERERTONAKSPTDADKAVITVSNMANGELLASNIMSESVAVALLTYKQPV 1267

```

Db 1062 NKSALKEAMPDDDDQFPDD-----LSLDHSD---LVAELLKEIS-----NHMERV 1104
 Qy 1268 ELLQTSLTNLGICIGKNCGLPNKHSRBMMLNLLEAEHTDVYIAGLHYLSKIMRSNK 1327
 Db 1105 EERKIALTEMLKLTQESBSFVWDEHFKTLLLTLDGDEPTIRALAKVLRELHQRP 1164
 Qy 1328 MRHMMHFLLELILKTIQCYOH-SKEALRDISMIRIAPSLPLDLISINIVPIATGEF 1386
 Db 1165 AR--FKYVAELTVKMLTELAHDPHKEVRSAEASVLATSIPEQCIXVLCPIIQTADY 1222
 Qy 1387 PTNLCAIKLLEVEHSGSETDAHLDIYFNLASADDTQSMKRAVFCVLKLYFTLG 1446
 Db 1223 PINILAIKQTKVIERVSKETLNLPLPEIMPGTIOGYDNSESVKACVCLVAVHAVIG 1282
 Qy 1447 EEKVKPKLSVLPNSKVRLLNVTIEKORNCISGGSGSTKNSAAS 1490
 Db 1283 DE-LKPHLSQLTSGSKMLNLYI-KRAQTSGGADPTTDVSGQS 1324

RESULT 6
 AAM39287
 ID AAM39287 standard; protein; 1295 AA.
 AC AAM39287;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 2432.
 XX
 KW Human; nocrotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYGE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QH;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 DR MPI: 2001-442253/47.
 DR N-PSDB; AAI58443.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders such
 XX as central nervous system injuries.
 PS Example 4; SEQ ID NO 2432; 10078bp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AAM38642-AAM4213) with nocrotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activites such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification

XX
 SQ Sequence 1295 AA;
 Query Match 17.7%; Score 1344.5; DB 4; Length 1295;
 Best Local Similarity 29.8%; Pred. No. 1.2e-91;
 Matches 404; Conservative 237; Mismatches 523; Indels 191; Gaps 40;

Qy 237 SALK-----NTNGNGVGLDEADNIGLRPRPTMIRKPLHSVSSLSLRPPNNV-DVTGDAG 291
 Db 31 SAFKVPAPKTSGN-----PANSARKP-----GSAGGFVAGASKEGAG 70
 Qy 292 AVTWSEFESSFEVVPOLNIFPAKMDIYQVLIISDKNADMEKRVDAKKIRALL-- 349
 Db 71 AVDEDDFIKAPTDVPSIQYSSRELEETINKIREILSDDKHDMQORANALKIRLLVAG 130
 Qy 350 -LSYHTQPFVAVQKELSLSFVDILKEELASOYIREACTITAVMSKTLRNKDLAFCSI 408
 Db 131 AAOYDCFPQHLRLDGLAKLS-----ADLNSQVREACTIVAHSLTVLGNKPFHGAGAI 185
 Qy 409 LEHLINLIONSAKYIASASTIALKYLKYTAAPKLLKYDTLNOGSKDIRSTLCENV 468
 Db 186 VPTLFNLVPSAKWATISGCAIRIRIRHTVPRILPITSNCT-SKVVPARRSFEFED 244
 Qy 469 LTFEEMQTKALERNATVLRDTLKSIGDADCARRHSFYAYAFRRHPELADQYGTLD 528
 Db 245 LLLQEMQHSLEHHAVALVETIKKGIHDADAEARVEAKTYGLNHLPEGEETLYNSLE 304
 Qy 529 IAAQALREREREGGGGGGTGTCTGTAPETRRTVSIRIGTPTLOKPTSMRISAVDPA 588
 Db 305 PSYQKSLQTYKSSG-----SVASLPQSDSSS---SQESLNPFPSSKMTANPSTVA 355
 Qy 589 ----AQRAKVAQVLYTSRQRPKPLGNPNNSNQASMTGAAA-----SGSLPRPRLNSNGGT- 639
 Db 356 GRVSAKSSKASSLPSSLQRSKSDIVNAAAGAKAHAAAGSVRSRIGALNAGSYASL 415
 Qy 640 -----PATTPGSVTPPRGRAG---VSQSPGSRSTP-----STKLADQYGIEN 682
 Db 416 GRVRAKLSAPLAGMNAADASGRSRTYKAVSQSGSRSGSGRYLTTTALSTVSSGVOR 475
 Qy 683 YVRGATGAI PKKASGIPNSTASSRETSPTR---SGGGLMKRMYSTGA---GSRRTPEHN 736
 Db 476 VLWNSASA--QKRSXIPRSGGCSREASPSRLSVARSSRPFRSVSGGCSREASRESRDT 533
 Qy 737 NVRPSAP-----ARLTAQSRHEAHTLGVGDGDP 766
 Db 534 SPVRSFQPLASRHSRSTGALYAPVYASGCGYIGSOSRSLSSSVASARVNTGSD--V 591
 Qy 767 DYVSGDYMSGGEMGRK-----LMGRDESDIDSEASVCSERSFPSSYTRGNKSNY 819
 Db 592 EBAVADALLGDIRTKKPARRRYESYGMHSDDDANSDASACSESRYSYS-----RNG 644
 Qy 820 SLGSHTRLDMSTORAPPDITIIQFCASHTMSERKDLISLTQVLAQCKELTQOOLKC 879
 Db 645 SI-----PTYMRQTEVAEVLNRCASSMSSEKBLGLQNLKXORTLSREBLK 695
 Qy 880 VLDMEKXEMDHTK-VSLFLDTVELILVANETSRNGSSCLTRLFNKLTDLNLSM 938
 Db 696 LCEIFTRMRADPHGRKVFEMFLETIVDFIQVHKDDL-QQMLFLLQLKKMGADLLGSV 754
 Qy 939 HSKWTKIQVHVEYPTQOLKELFRILISDSTQPTTKRIARLRLTLANTYCKSSDF 998
 Db 755 QAKVQALVNTESFPNDQFNILMRFTVDQOTBELKVAIKVAILKYEITLA-----KQMD- 809


```

QY 999 PSD--OSQACERTVYLKLAQLADQKSMELRSQASCLVALYNLTPOMTLLADLPKYQ 1056
D 810 PGDEINSEETRLAASRVITMTTEPKSSVPRKAQSVLISLELTPEFTMLGALPTFQ 869
QY 1057 DSAASCISHHRRRSQSCNSGANSPPSSPLSSSPK-----PLOSPPVCPFASLOSH 1109
D 870 DGAATKLLHNHLR-----NTGNGTQSMGSPLTRPPRSPANSSPLTSPNTSQNTLSPSA 925
QY 1110 HQLSISSTSPRSSVSE-----QELFSSELDIQHNIQKTSSE---IRHCFGGQYQTA 1160
D 926 FDVTEEMNSEDIYSILRGVTEALQNFSPQEDMEELKDKDKDGDGSCG-----PG 981
QY 1161 LAPNGFNHLYQHDOGQSCASLSSNSKTQSSANTQSN---TPESATRLDNLERER 1216
D 982 MSDPRAGDADTDSQTALDNKASLLHSPHTSSPRSDYNNYNSDSISPNNKALKEAM 1041
QY 1217 TTQNAKSPPTDAKITYSINNAENGELLASNLMESEVVRVALTLTDQPELLQTSLTN 1276
D 1042 FDDADQPPD-----LSLDHSD-----LVAELELS-----NHERVEERKIALYE 1084
QY 1277 LGLCIKGCNCELPNKHFRSIRMLNILEAHTDVVIALGLHLSKINRSMKRNHMHFL 1336
D 1085 LMKLTQESPSVMEHETKTLILLLETLGDEPTIRALALKVLEIRHQPAR--FNQYA 1142
QY 1337 ELILKTIQCYQH-SKALRDIDSMIRIAPSLPLDSINIVNPAITGEPTNLCAIKI 1395
D 1143 ELTVKTLLEAHKDPHEKVSASAEBAASVLTATSIPECQIKVLCPIIQTADYPIMLAIKM 1202
QY 1396 LLEVTENHGEITDAHLDIYEPNLABADDTQSNVRAAFVCIYKLYFVGEENVRKLS 1455
D 1203 QTKYIEVSKETLLNLLPEIMPGLIQGVNSESSEVRACVCLVAHVAIGDE-LKPHLS 1261
QY 1456 VLNPSKRYLLNVLEYKORNCISGGSSSTKNSSAAS 1490
D 1262 QLTGSKKMLNLYI-KRAQTGSGGADPTTVSGGS 1295

```

```

PA (HYSE-) HYSEQ INC.
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI, 2001-442253/47.
DR N-PSDB; AA158442.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
PS
PS Example 4; SEQ ID NO 2431; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AA38642-AA42213) with neurotrophic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 1309 AA;
QY
QY Query Match 17.7%; Score 1344.5; DB 4; Length 1309;
QY Best Local Similarity 29.5%; Pred. No. 1,2e-91;
QY Matches 405; Conservative 233; Mismatches 523; Indels 211; Gaps 40;
D 237 SALK-----NTNGNGVGLDEADNIGLRERPTMKRPLHSAVSSSLRPKNVN-DVYGDAG 221
D 31 SAKVPAKPTSGN-----PANSARKP-----GSAGGPKVAGASKEGAG 70
QY 292 AATMESFESSFEVVPOLNIFHAKMDMDIYKOVLVITSDKNADBEKVDALKIRALLI-- 349
D 71 AYDEDDPFAFTDVBEOIYYSRELEPTLANKIRELSDKHMDQAMLKIRISLLVAG 130
QY 350 -LSYHTQPOFVAVOLKELSLSPVDILKEELRSQVIREACTITTAWSKTLRNLDAFCWSI 408
D 131 AAQYDCFPQHLRLDQALKLS-----AKDLRSQVVEACITVAHLSTVLGNKPDHGAELI 185
QY 409 LEHLTLVLIONSAAKVIASASTIALKTIYIKTHAPKLIKITYDTLNOKSDIRTLCELVW 468
D 186 VPTLFWLVNSAKVWATSCAARFIRTHVPRLLPLTISNCT-SKSVVPRRSFEFLD 244
QY 469 LTFEEQOTALARNATVLDLTLLKSGIDGACARRSRRAVYAFRHPFELADQIYGTLD 528
D 245 LLLQEWQTHSLERHAAVAVETIKKGIHDAERARVAKTYMGLRHFQGEAETLYNSLE 304
QY 529 IAAQALEREREGGG-----GGTGTGTAPETRTVSR 563
D 305 PSYQKSLQTYLKSSGSVASLPQSDRSSSSSQSLNRPFSSKWTANPSTVAGVSAAGSK 364
QY 564 IGRTPGTLOKPTPSMSISAVPTAAQAKVPAQVLYLRQKKPLGPNNNSQAS----- 617
D 365 ASLPGSLQ-----RSRSDIDVNAAGAKAHAAQAQSVRSGR-LGAGALNAGSYASLED 417
QY 618 ----MTGAASGLLPRPLRNSNSGTPATTPSVPRPRGRAG---VSOQOPGSRSTSP- 669
D 418 TSDKLDGTASBDGRRAKLSA-----PLAGMNAKADSGRRTKRVVSSQPSRSGSGC 422
QY 670 ----STKLDQYGGIGNYRGATGAI PKKASGIPSTASRSTSPTR---SGGGLKRSN 722
D 473 RVLTTLTALSTVSSGVORVLVNSASA--QKRSKIPRSGGSRBASPSRLLVASSSRIPRS 530
QY 723 YSTGA---GSRTTPERNNVPSAP-----ARLLA 749

```

```

Db      531 VSGGCSREASRBSRDTSFVRSPFOPLASRHHRSSTGALVAPRYGASGFGYIGSSSRIS 590
Qy      750 QSRREHLTGVDGDDQDPDVVSGDVMRSGGMRGRK-----LMGRDESDDIDSEASSVC 802
Db      591 SSVASMRVNTGSD--VEEAVADALLGDIRTKKPPARRRYSYGMHSDDDANSPASSAC 648
Qy      803 SRSRFSSTYTRGNKSNYSLSGSHTRLDWSTORAPFDIETIIIOFCASTHMERKQGLISL 862
Db      649 SRSRSTSS-----RNGSI-----PTVMRQEDVAEVLNRCASSWSRKGGLGL 692
Qy      863 TOYLADKELTOQOLKCVLMFRKMPDTHYK-VISLFDITYTELLVHANTSRNGSSS 921
Db      693 QNLKNORTLSRVELKELCEIFTRMPADHGKRVESFMELETVDFIOVHKDL-ODWLPV 751
Qy      922 CLTRFLNKLGTLLNSMHSKIKITLOVHVEFPYTOLELRIISDSQPTTKTRIAI 981
Db      752 LITOLLKKGADLLGSVQAKVQALDVTRESFPNDIQENILKRFVVDOTQTSKAKVAI 811
Qy      982 LRFITDLANTYCKSDPSPD--QSQACERTVLKLAQLADOKSMELRSQARCLVALYNL 1039
Db      812 LKRIETLA-----KQMD-PGDFINSSETRLAVSRVITWTEPKSSDVKKAQSVLISLFE 866
Qy      1040 NTPQMTLLIADI.PKYVQDSARSCISHMRQSQSCNSGANSFSSSPLSSSK----- 1092
Db      867 NTFEFTMLIGALPKTFQDQATKLIHNLH-----NTNGTQSSWGSPLETPRPSANWSS 922
Qy      1093 PLQSPSVGFPALQSHHHLQSLSSSTSPRROSSVE-----OELPSESLDIQHNIOKTS 1146
Db      923 PLTSPNTISQNTLSPSAFDYDTENNMSEDIYSRLGVTEALIQFSPRQEDWMEPLKRS 982
Qy      1147 EE---IRHCFGGQYOTALPNGFNGLQYHDOQODSCASLSNSKTQSSANTTQSN--- 1200
Db      983 KQDDGDSMCGG---PGMSDPRAGGATSSQOTALNKSLSLHSMPTSSPSRDNPNYN 1038
Qy      1201 -TPESATMBLMDLERERTQNAKSPTDAAKIVTSSINMANGELILASULMESEVVRVAL 1259
Db      1039 YSDSISPFKSAKLEAMFDDADQFPD-----LSLDHSD---LVAEELKELHS----- 1083
Qy      1260 TLTKQOPVELLQSLNLGICIGNGCELPNKHFRIMMNLINLAETHVDVITAGIAYL 1319
Db      1084 --NNHNRVERKALVLEMLKLTQESFSFVDEHFKTLITLLTLDGKEPTIALAKVL 1141
Qy      1320 SKIMRSNKRHHNMHLELILKIIQCYOH-SKEALRIDISMPRIAPSLPLDSINIVN 1378
Db      1142 RETLRHQPAR--FKNYAEELVKTLEHNDPKHEVRSABEASVLAATSPQCICKVLC 1199
Qy      1379 PVATGEFPTNLCAIKILEVTEHSGSEITDAHLDIVFPNLARSADDTQSMVKAAPCT 1438
Db      1200 PIQTADYPIINLAIMQTFVIERVSKETLNLPLPEIMPGLIQGYDNSESSYKACVFCI 1259
Qy      1439 VKLYFLVGEERKPKLSVLPKSVRLNIVTEKORNCISGGSGSTKSSAAS 1490
Db      1260 VAVHAIVGDE-LKPHLSQLTSGSKKCLNLYI-KRAQTSGGADPTTIDVSGQS 1309

```

RESULT 8
AAM41072
ID AAM41072 standard; protein; 1317 AA.

AC AAM41072;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 6003.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.

```

OS Homo sapiens.
XX WO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US034263.
XX 23-DEC-1999; 99US-00471275.
XX 21-JAN-2000; 2000US-0048725.
XX 25-APR-2000; 2000US-00552317.
XX 20-JUN-2000; 2000US-00598042.
XX 19-JUL-2000; 2000US-00620312.
XX 03-AUG-2000; 2000US-00653450.
XX 14-SEP-2000; 2000US-00662191.
XX 19-OCT-2000; 2000US-00693036.
XX 29-NOV-2000; 2000US-00727344.
XX (HUSE-) HUSE INC.
XX Tang YF, Liu C, Aundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX Zhou P, Goodrich R, Drmanac RT;
XX MPI: 2001-442253/47.
XX N-PSDB; AAI60228.
XX Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
XX Example 2; SEQ ID NO 6003; 1007bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
XX encoded polypeptides (AAM36642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemia and
XX C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
XX
XX SQ Sequence 1317 AA;
XX
XX Query Match 17.7%; Score 1344.5; DB 4; Length 1317;
XX Best local Similarity 29.5%; Pred. No. 1.3e-91;
XX Matches 405; Conservative 233; Mismatches 523; Indels 211; Gaps 40;
Qy 237 SALK-----NTNGNGUGLDADNIGLRERPTMKRPLHSAVSSLSLPPKNVN-DVTGDAG 291
Db 39 SAKVPAPKTSGN-----PANARKP-----GSAGPXYGAGSKSGAG 78
Qy 292 AVTMESEFSEFVVDQNLIFHAKMDIYQVLIISDNADWEKRVDAIKTRALLI-- 349
Db 79 AYDEDDFIKAFIDVPSIQIYSRELEETLNKIRELISDDKHMDQDANALKTRISLVAG 138
Qy 350 -LSYHTQPOFAVQKELSLSFVDILKELEBQVIRACITTAAYMSKTLRNKLDACWSI 408
Db 139 AAQYDCEFOHLRLDLGALKLS-----AKLRSQVAREACTIVAAHLSLVGNKFDHGAELI 193
Qy 409 LEHLINLQNSAKVINGASTIAKLYIKYTHAPKLIKITYDTLQNSKXDIRSTCELVY 468
Db 194 VPLFLVFNASKAKWATSGCAIRIIRHTHPRLIPLTNSCT-SKSVVRRSRSEFLD 252
Qy 469 LTFEEMQTKALERNATVLDTLKSGSDGADCDARRHSRYAYVAFRRHPPELADQITGLD 528
Db 253 LILQEWQHTSLERHAVALVETIKKGIHDADAEARVARKTYMGLRNHFPGEAEITLVNSLE 312

```

```

QY 529 IAAQALREREREGGG-----GCTGTGTGTAPETRTVSR 563
DB 313 PSYOKSLQTYLKSSGSVASLPQSDRSSSSQESLNRPFSSKMTANPSTVAGRVASGSK 372
QY 564 IGRFTGTQKTPPSKRSISAVDTAAQAKVRAQTYLTSRQKPLGPNNSQAS-----617
DB 373 ASSLPFGSLQ-----RSRSDIDVNAAGAKAHHAAGQSVRSGR-LGAGALVAGSVASLED 425
QY 618 ----MTGAASGSLRPRPLNSNGCTPATTPGCVPRPRGRAG---VSQSPGSRSTSP- 669
DB 426 TSDKLDGASDEGRRAKLSA-----PLAGMNAKADSRGSRMTMSQSPGSRSGSPG 480
QY 670 ----STKLRDYGGIGNYRGATGAIPKASGIPRSTASRETSPTR---SGGLMKRSM 722
DB 481 RVLTTLTALSTVSGVQVLVNSASA--QKRSKIPRSGCSSEAPSRSLSTVARSRIRPPS 538
QY 723 YSTGA---GSRTPERRNNPVRPSAP-----ARLLA 749
DB 539 VSQCSREASRESRSDTSPVRSFQPLASRHHSRGTALYAPEVYGAAGPGYGISQSSRLS 598
QY 750 QSRBAHTLGVGDGQDPYVSGDYMRSGGMGRK-----LMGRDESDIDSEASVC 802
DB 599 SSVAMRVLTNGSD--VEEAVADALLGDIRTKKPPARRRYESYGMHSDDANSASAC 656
QY 803 SERFSDSYTRGNKSNYSLSGSHTRLDWSTORAFDIIETIICFCASTHSEKDGILSL 862
DB 657 SERYSYSS-----RNGSI-----PTMRQTEDEVAELNRCASNMEREGGLGL 700
QY 863 TQYIADGKELTQOOLKCVLDMFRKKFMDTHK-VYSLFLDTVTLELLVHANETRNGSSS 921
DB 701 QNLKNQRTSLRVELKRLCEIFTRMFADPHGRKVFMSFLETLVDFIQVHKDDL-QDMLFV 759
QY 922 CLTEFLNKGTDLNSMSKIMKTLQVHEXEPQLOKELPRIISDSTOTPTKTHAI 981
DB 760 LLTQLKKMGADLLGSVQAKVQKALDVRESFPDLQNLIMRTVDOTQSLKVKAI 819
QY 982 LRFLTLANTYCKSSDPPSD--OSQACERTVLKLAQLAADQSMELNSQARSCVALVYL 1039
DB 820 LKYEITLA---KQMD-PGDFINSETRLAVSRVITWTEPKSDVRAAQSVLISLFL 874
QY 1040 NTPQMTLLADLPVYQDSARSCIHSHMRROSQSGNSGANSPLSSSPK-----1092
DB 875 NTPFTMLLGLPFTFDGATKLLHNLH---NTGNGTQSSMCSPLRPPRPSANWSS 930
QY 1093 PLOGSPVPFASLQSHHQLSISSTSPRSQSV-----QELLFSEELIQHNIOKTS 1146
DB 931 PLTSPFTNTSONTLSPSADFDTENMNSEDIYSSLRGVTEALQNFSPSQEDMNEPLKRS 990
QY 1147 EE---IRHCFGQYQOTALPANGFNGLQYHDQGOODSCASLSNSKTQSSANTTQSN-- 1200
DB 991 KKDGDGSMCGG---PGMSDPRAGDDATDSSQTLADNKASLLHMSPTTSSPRSDYPN 1046
QY 1201 -TPSSATMRDLNLERERTQAKSFPTDQAKVITVSINANGELILASNLMESEVRYAL 1259
DB 1047 YSDSISPNKSALEAMFDDADQFPDD-----SLDHS---LVALLKELS-----1091
QY 1260 TLTDQDPELLQTSITNLGICIKGNCCLPNKFRSIRMNLNLEHTDVTYAGLHV 1319
DB 1092 --NHNERYEERKIALYELMKLTQEEBSFSVWDEHFTLILLLETTGDEPTRLALVKL 1149
QY 1320 SKIRSNMRRNMHMFLELLIKLIIQCYOH-SKEALPDIDSMIPRIAPSLPLDISINIVN 1378
DB 1150 REILRHQAR--FNKYVELTYMKTLEAHKDPHKEVVRGAESAVALATISISPECIKVLC 1267
QY 1379 PVIAITGEPTNLCAIKILEVTEHHGSEITDAHLDYFNPULARSADTQSNVRAAFCT 1438
DB 1208 PIITADYPIIMAIKQTKVIERYSKETLNLILPEIMPGIIOGDNSESSVRAACVFC 1267
QY 1439 VKLFFVGEKVKRSLVNSKVLNLYEKOQNCISGGGSSSTKXSSAAS 1490
DB 1268 VAVHAYIGDE-LKPHLSQLTGSKKMLNLVYI-KQAQTSGGADPTTDSGOS 1317

```

```

RESULT 9
AAM41073
ID AAM41073 standard; protein; 1317 AA.
XX
AC AAM41073;
DT 22-OCT-2001 (first entry)
DE Human polypeptide SEQ ID NO 6004.
XX
KW Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000MO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUL-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac R;
XX
DR WPI; 2001-442253/47.
XX
DR N-PSDB; AAI60229.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
as central nervous system injuries.
XX
PS Example 2; SEQ ID NO 6004; 10078bp; English.
XX
The invention relates to human nucleic acids (AA157798-AA161369) and the
encoded polypeptides (AAM38642-AAM42213) with nocotropic,
XX
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
in gene therapy. A composition containing a polypeptide or polynucleotide
XX
CC system, such as peripheral nervous injuries, peripheral neuropathy and
centralised neuropathies and central nervous system diseases, such as
XX
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX
CC utilisation of the activities such as: Immune system suppression,
activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic
XX
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
assays for receptor activity, arthritis and inflammation, leukaemias and
XX
CC C.N.S disorders. Note: The sequence data for this patent did not form
part of the printed specification
XX
SQ Sequence 1317 AA;
XX
Query Match 17.7%; Score 1344.5; DB 4; Length 1317;
Best Local Similarity 29.5%; Pred. No. 1,3e-91;
Matches 405; Conservative 233; Mismatches 523; Indels 211; Gaps 40;
XX
237 SALK---NTNGNGVGLDEADNIGLREPTRMKIRPLHSAVSSSLRPKRVN-DVTGDAG 291

```

```

Db      39 SAFKVPAPKTSGN-----PANSARKP-----GSAGGPKVAGAGSKXGGG 78
Qy      292 ATMTSEFSEFVVPOLNIFHAKMDIYKQVLIISDNADBEKXVDLKKIRALLI-- 349
Db      79 AVEDEDFIAFDVPSIQIYSSRELEETLNKIRELISDDGHDQDQANLKKIRSLVVG 138
Qy      350 -LSYHQPOQFVAVOLKELSLSPVDILKEELRSQVIREACTIIAYMSKTLRNKLDAFCWSI 408
Db      139 AAQYDFFQHLRLDGLAKLS-----AKDLRSQVREACTIVAHLSVLGNKFDHGAEMI 193
Qy      409 LEHLINLONSXAVIASASTIAIKYIKTHAPKLIKITDTLNOSKSDISTCELV 468
Db      194 VPTFLVFNPSAKVMATSGCAIRFIIRHVPRLPLITSNCT-SKSPVRRSRPEFLD 252
Qy      469 LFEEMQTKALENNAVLVDLTKKSGDGDCCDARRSRAYAYAFRRHPELDQIYGTLD 528
Db      253 LLLQEMQTHSLERHAAVLVETIKKGIHDADAEARVEARKTYMGLRNHPGEAETLYNSLE 312
Qy      529 IAAQRLEREREGGG-----CGTGCTGTAPETRRTVSR 563
Db      313 PSYQKSLQTYLKSSGVASLPQSDRSSSSQESLNRFSSKMTANPSTVAGRVASGSK 372
Qy      564 IGRTPGLQKPTPSMKSISAVDTAAQRAKRAQYTLYSRQKPLCPNNSNQAS----- 617
Db      373 ASGLPGSLQ-----RSRDIDVNAAGAKAHHAAGSVSGR-LGAGALNAGSVASLED 425
Qy      618 ----MTGAAASGSLPRRLNSNGGTPATTPGVSITRRPRRAG---VSSQPSGRSTSP- 669
Db      426 TSDKLDGTASDEDRVAKLSA-----PLAGMNAKADSRSRSTKWSQSPQSRSGSPG 480
Qy      670 ----STRKLDQYGGIGNYRGATGALPKKASGIPRSTASSREHSPTP---SGGGLMKRSM 722
Db      481 RULTTALSLVSSGVQRAVLVMSASA--QKRSKIPROGCGREASPSRLSVASRSRIPRPS 538
Qy      723 YSTGA---GSRRTPERNNPVRPSAP-----ARLLA 749
Db      539 VSGGCGREASRESRSDTSPVRSFQPLASRHSRSTGALVAPEYVAGSGPGYGISQSRSL 598
Qy      750 QSRFAEHTLGVGDGQPDVYSGDYMRSGMRMGRK-----LMGRSDSDIDSEASVVC 802
Db      599 SSVSARKVAVMTGSD--VEEAVADALLGDIRTKKKPPARRRYESYGMHSDDANSASSAC 656
Qy      803 SESSPSSYTRGNKSNVSLSGSHRLDWSIORAPFDIETIIIOFCASTHWSERKGLISL 862
Db      657 SESSYSS-----RNGSI-----PYTMQTEVAVALNRCHASSMSEKKEGLGL 700
Qy      863 TOYLADKEITLQOOLKCVLDMFRMFMNDHTK--VYSLFDVTVELILVHANETSRRNGSS 921
Db      701 QNLKNGRTLSRYELKRLCEIFTRMFADPHGKRVFSMFLFVDFIQVHKDDL-QDWLFV 759
Qy      922 CLTRLFNKGLTDLINSNASKIKWTLQVHVHFFPQLQKELFRISISTOTPTTKRIAI 981
Db      760 ILTQLLKKMGADILGSAQAKVQKALDVTRESFPDLOFNILMRTVYQOTPSLKVAVAI 819
Qy      982 LRFUTDLANTYCKSSDPSD--QSQACERTVLKLAQLAADQSKMELRSQARSCVAVLYNL 1039
Db      820 LKYLETIA---KQMD-PGDFINSETRLAVSRIYTTTEKSDVYKAAQSVLISLPEL 874
Qy      1040 NTPQMTLLADLPKYQDSARSCISHMRROSCNSGANSPPSSPLSSSPK----- 1092
Db      875 NTPFTMLLALPTFPDGAATKLHNHR-----NTGNGTQSMQSPLTRPFRPAMWSS 930
Qy      1093 PLOGSPVGPASLGSHHQLSISTSPRSRQSSVE-----QELFSEELDIGNIKTS 1146
Db      931 PLTSPITVTSQNTLSPSAFDYDTEMNSEDYVSSLRGVEALQNFSPSQEDWNEPLKSDS 990
Qy      1147 EE---IHRCGGQYQTLAPGFGNGLQYHDQGDSCASLSNKSQSSANTQSN--- 1200
Db      991 KKDDGDSMGCG---PGMSDPRAGDADTDSQTLADNKASLILHMPHTSSRSRDYVPYN 1046
Qy      1201 -TPESATWRLDNLERERTONAKSPTDPAKVIIVSINNAENGELLASNLMESEVAVAL 1259

```

```

Db      1047 YSDSISPPNKSALEKAMPDDADQPD-----LSLDHSD-----LVAEELKELS----- 1091
Qy      1260 TLTQDQVELLIQTSITNIGTICKGNCELPNKRPISIRMLNLTLEAHTVDVINGHLV 1319
Db      1092 --NHNREVEREKIALYELEMLKLTQESFSWDEHFTILLTLLLETGDDKEPTRALAKVL 1149
Qy      1320 SKIMSNMGRNHNMFLELILKTIQCYOH--SKELARDIDSMIPRIASLPDLISINIVN 1378
Db      1150 REILHQPAR--FKNYAEITWTKTEIAKDPKKEVYRGAEEASVLAIVLSISPEQCIKVL 1207
Qy      1379 PVATGEPTMLCAIKILLEVTEHNSBITDAHLDIVFPNLRASADDTQSMYKAAVFCI 1438
Db      1208 PIIQIADPFINAIIKMQTKVIERVSKETLNLILPEIMGGLIQGYDNSSSSVKACVFL 1267
Qy      1439 VKLYVLGEKVKPKLVSANPSKATLVNVIKQNCISGGSSSTKNSAAS 1490
Db      1268 VAVHAVIGDE-LKPHLSQLTGSKMGLNLNLYI-KRAQTSGGADPTTVSGQS 1317

RESULT 10
ADB64143
ID      ADB64143 standard; protein; 916 AA.
XX
AC      ADB64143;
XX
DT      04-DEC-2003 (first entry)
XX
DE      Human protein encoded by clone CTONG20032930.
XX
KW      Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
KW      cell regeneration; membrane protein; signal transduction-related protein;
KW      transcription-related protein; osteoporosis; neurological disease;
KW      cancer; tumour.
XX
OS      Homo sapiens.
XX
PN      EP1308459-A2.
XX
PD      07-MAY-2003.
XX
PF      28-MAR-2002; 2002EP-00007401.
XX
PR      05-NOV-2001; 2001JP-00379298.
XX
PR      25-JAN-2002; 2002US-00350978.
XX
PA      (HELI-) HELIX RES INST.
XX
PA      (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI      Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI      Yamamoto Y, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI      Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuno Y;
XX
WP1: 2003-450961/43.
XX
PS      N-PSDB; ADB62173.
XX
PT      New polynucleotides and polypeptides, useful for developing a diagnostic
PT      marker or medicines for regulation of their expression and activity, or
PT      as targets of gene therapy.
XX
PS      Claim 1; Page; 222pp; English.
XX
CC      The invention discloses a polynucleotide comprising a sequence selected
CC      from 1970 fully defined nucleotide sequences which encode novel
CC      polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC      or its partial peptide, an antibody binding to the polypeptide or peptide
CC      of the polynucleotide, immunologically assaying the polypeptide or peptide
CC      peptide of the polynucleotide by contacting the polypeptide or peptide
CC      with the antibody of the encoded protein, and observing the binding
CC      between the two, a transformant carrying the polynucleotide in an
CC      expressible manner and an antisense polynucleotide. The oligonucleotide
CC      is useful as a primer for synthesising the polynucleotide, or as a probe
CC      for detecting the polynucleotide. The polynucleotides and encoded
CC      proteins are useful as pharmaceutical agents and many disease-related

```


transcription-related proteins, disease-related proteins and genes encoding them can be used as indicators for diseases (e.g. osteoporosis, neurological diseases, cancer, tumours. The cDNA may be used to regulate the activity or expression of the encoded protein to treat diseases. The sequence presented is a protein of the invention. Note: Some of the sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office.

Sequence 723 AA.

Query Match 12.1%; Score 920; DB 7; Length 723;

Best Local Similarity 32.0%; Pred. No. 6e-60; Matches 243; Conservative 144; Mismatches 274; Indels 98; Gaps 19;

5 KPESDLOFIOMPRADNRKVOLAEDLVTEL-SDDTNSICTDMDGFL---IDGIMPVLTG 60-
2 EPRSMETFCAGVOQKDGRLQVQGLVYAGALSDLEBDGRLGKIVDALTGWVS 61
61 SHFKIAQKSLAEFSELIKRLGSDENATATVLPVHIDRLGDSRDTVEKAQLLRDMEH 120
62 SNRVSLMGLEILSAFDRILSTRKSYAAVIVALIDRMGDAXKVRDEAQTILKLM-D 120
121 RVLPQALIDKATSCFRHKNKAVREEFLOTIVNAHEGTQQLSVAVYIPVCAIGDP 180
121 QVAPPMYIWEQLA-SGFEKHKVFRSEGVCLLETLLNIFGAQPLVISEKLIPLHLCILFSDS 179
181 TAVNREAIQTIVYVGVGDRLARPDLRRMDVPASLAMEQFEDVQKQGLLPPALK 240
180 NSQVRDAIILAIVETIYHVEGKMDLYKR-GIPPALEWIFAFEDVQSSGGMILSVCK 238
241 NTNGNGVGLDEADNIGRERPT-----RMIRPLHSAYSSSLRPKPNVNDYT- 287
239 DKS-----PDDEEVD-GNRPSSAASAFKVPAPRTSGNPNANSAPKPSAGSPKVGASKE 292
288 GDAGAVTMSFESSFEVVPQINIFHAMDMDIYQVLIISDKNADNEKRVDAIKTRAL 347
293 GGAGAVDEDDPIKFTDVPISIQIYSSRELETLNKIREILSDGDHWDQANALKKIRSL 352
348 LI---LSYHPOPOVAVQKELSLSPFDILKEELRSQVIRACTIYVMSKTLBNKLDAF 404
353 LVAGAAQYDGFQHLRLDGLALKLS-----AKDLRSQVREACTIVHLSLVLANKKFDHG 407
405 CWSLIEHLINLIQNSAKIVASASTIALKYIIKYTHAPKLEKIYTDLTNOSKXDIRSTLC 464
408 ABAIVPTLPHLVPSAKVATSGCAALRFRIRHVRRLIPLITSNCT-SKSVVRRRSF 466
465 ELMWLFEEMQTKALERNATVLRDLTKSSIGDADCDARRHSRYVYMAFRHHPPELADQIY 524
467 EPLDLLLQEWQTHLEHAAVLVETIKGIHDADAARVARKTYMGLRNHPGAELETLY 526
525 GTLDIAAQRALERRRGGG-----CGTGTCTGAPETRR 559
527 NSLIEPSYOKSLQTYLNSGVSVALPSQDRSSSSSSQESLNRPFSSKMSWTANSTYAGRVA 586
560 TVNSIRGPTGLQKPTSMRSISAVDPAARAKVRAQYLYTSRQKPLGPNNSQMS-- 617
587 GSSKASSLPESLQ-----KRSRSDIVAAAGAKAHAAQGSVSGR-LGAGALNASSYA 639
618 -----MTGAAASGSLPRRLNSNSGCTPATTPGVSVPFRPAG--VSQSQPGSRS 666
640 SLEDTSDLDGTASBEDGRVAKLSA-----PLACGKGNKAKDSRRTKWSQSQPGSRS 694
667 TSPETKLADYGGIGNYRGATGALPKKASGIPSTASS 705
695 GSPGRVLT-----TTTALSTVSSGVQRLVNS 720

RESULT 12
ADU20651
ID ADU20651 standard; protein, 1330 AA.
XX
AC ADU20651;

13-JAN-2005 (first entry)
A. thaliana At2g20190 homologue.
plant; tolerance; resistance; environmental stress; drought.
Arabidopsis.
thaliana.
MO2004092349-A2.
28-OCT-2004.
15-APR-2004; 2004WO-US011887.
15-APR-2003; 2003EP-00008079.
01-AUG-2003; 2003EP-00016671.
30-SEP-2003; 2003EP-00022226.
(BADI) BASF PLANT SCI GMBH.
Puzio P, Chardonnens A, Chen R, Puente P;
WPI, 2004-766856/75.
N-PSDB; ADU20650.
New transformed plant cell with altered metabolic activity compared to a corresponding non transformed wild type plant cell, useful for producing, screening and breeding plants with increased tolerance to environmental stresses.
Disclosure; Page 598-599; 607pp; English.
This invention describes a method resulting in a novel transformed plant cell which has been altered by an inactivated or down-regulated gene and results in increased tolerance and/or resistance to an environmental stress as compared to a corresponding non-transformed wild type plant cell. The method can be used to generate or screen for seeds or plants with increased tolerance/resistance to environmental stress and for detecting stress in cells or plants. The transformed plant cell comprises a metabolic activity that is altered concerning one or more metabolites selected from 2,3-dimethyl-5-phytylquinol, 2-hydroxy-palmitic acid, 3,4-dihydroxyphenylalanine, 3-hydroxy-palmitic acid, 5-oxoproline, alanine, alpha linolenic acid, alpha-tocopherol, aminoacetic acid, abhydroglucose, arginine, aspartic acid, beta-apo-81 carotenal, serine, beta-carotene, beta-sitosterol, beta-tocopherol, hexadecatrienic acid, margarinic acid, gamma-aminobutyric acid, (delta-7-cis,10-cis)-hexadecadienic acid, delta-15-cis-tetracosenic acid, ferulic acid, campesterol, cerotic acid (c26:0), citrulline, cryptoxanthine, eicosenoic acid (20:1), fructose, fumarate, galactose, gamma-tocopherol, gluconic acid, glucose, glutamic acid, glutamine, glycerate, glyceraldehyde, glycerol, glycerol-3-phosphate, glycine, homoserine, inositol, isoleucine, isomaltose, isopentenyl pyrophosphate, leucine, lignoceric acid (c24:0), linoleic acid (c18:2 (c9, c12)), leucine, lycopene, malate, mannose, methionine, methylgalactofuranoside, methylgalactopyranoside, methylgalactopyranoside, palmitic acid (c16:0), phenylalanine, phosphate, proline, putrescine, pyruvate, raffinose, ribonic acid, shikimate, sinapine acid, stearic acid (c18:0), succinate, sucrose, threonine, triacontenoic acid, tryptophane, tyrosine, ubiquinone, udp-glucose, valine, and zeaxanthine. The plant is selected from maize, wheat, rye, oat, triticale, rice, barley, soybean, peanut, cotton, rapeseed, canola, manihot, pepper, sunflower, flax, borage, safflower, linseed, primrose, rapeseed, turnip rape, teagates, solanaceous plants, potato, tobacco, eggplant, tomato, Vicia species, pea, alfalfa, coffee, cacao, tea, Salix species, oil palm, coconut, perennial grass, forage crops and Arabidopsis thaliana. The environmental stress is selected from salinity, drought, temperature, metal, chemical, pathogenic and oxidative stresses, or their combinations. The inactivation or down-regulation of the gene is achieved by double-stranded RNA interference (dsRNAi), introduction of an antisense nucleic acid, a ribozyme, an antisense nucleic acid combined with a ribozyme, a nucleic acid encoding a co-suppressor, a nucleic acid encoding a dominant negative protein, DNA- or RNA- or protein-binding

CC factors targeting said gene or -RNA or -proteins, RNA degradation
 CC inducing viral nucleic acids and expression systems, systems for inducing
 CC a homologous recombination of the genes, mutations in the genes or their
 CC combinations. This sequence represents a homologue of an Arabidopsis
 CC thaliana drought tolerance protein.
 XX

Sequence 1330 AA;

Query Match 8.0%; Score 608.5; DB 8; Length 1330;

Best Local Similarity 20.1%; Pred. No. 5.6e-36;

Matches 295; Conservative 290; Mismatches 618; Indels 267; Gaps 52;

122 VLPQALDKATSCCKKNAKREBFLOTIVNALHEVGTQOLSV-RVYIPVYCALLGDP 180
 3 VSSPIIVERASVYMMHSMRVRREFPARTVSAIGLFSTELPQORVILAIPLMDLP 62
 181 TNNVEAIIQTLVEIKVYKVDRLPDLRMDVPASKMLQKPDQVQESLILPSALK 240
 63 NOAVEAALICEEMMGSGQFRELQK-HHLPSTWMDINARLERLEPQ-----LR 114
 241 NTNGNGVGLDEADNIGLRERPTMIKRPILHSAVSSSLRPKN-----VNDVTGDAG 291
 115 STDG-----RSAHVYVYKASVAVPKKSPAKAPATRENSLFGDA 156
 232 AVTMSESSFEVVDQNIIFAKDM-DDIYKQVLIISDKNADWEKRVDAIKIRALLIL 350
 157 DITEKPIE-----PIKYSEKELIREPEKIAATLVEPK-DWMSRISAMRVEG-LVA 206
 351 SVHTQPPAVQKELSLSPVILKEEL-RSVDIREFACTIATMSTLRKLDAPCMS 407
 207 GGATVYSCFGLLKQI---VGPLSTQLADRRSTVYKQACHLLCLLSKELLDFEACET 262
 408 ILEHLINILIONSAKYIASASTIALKYIIKYTHAPKLKIYDTLNQSKSDIRSLCELM 467
 263 FIPVFLKVLITVLAIESADNCIKTMLNCKAAVLPRIASAGHDNALIRACCEYA 322
 468 VLLFEEM-OTKALENNATVLRDTLKKSIGDADCDARRHSRYVYAFRRHPELADQIYGT 526
 323 LITLHEWDPAPFIQRSVDLYEDLIRCCVADAMESEVATFAMCYRMFAKTWPDERSRLSS 382
 527 LDIAORALERBEGGGGTGTGTGTAETRTYRIGRTGTLQKPTPSMRISAVDT 586
 383 FDPVITQRLINE---DGIHRRHASPVYRERHSQPSQTSAPSNLPEYGTSAIYANDR 438
 587 AAACRAKYAQTLYSRQKPLGPNNNSN-----QASMTGAASGLPR-----P 630
 439 SSNLSGSGSLSGLLSOSKDVNKGSESLSVLSSKQKSAISMRLGHLISDRNPA 498
 631 RLSNSGGTPTTPGSPVTPRPRGAGVSGSQPS---RSTSPSTLKDQYGGIGNYRGA 687
 499 ALRSSSLDLGVDPSPSRDPPFAVAVPASNSHTSSAAESTHSINKSNRNGGLG----- 552
 688 TGAIRKXASGIPRSTASSTRETSPTRSOGGLMKRMYNS-TGASRRTPRNNNVRSAPAR 746
 553 -----LSDIITQIQASKDSGRSSYKGNLSESHPTFSSLTAKKGSERNE-RSS----- 599
 747 LLAQREAEHT---LGVDDGQDPYVSGD--YMSGGGMM-----GRKLNG 787
 600 -LEESNDAREVRFMAAGHDRQMDTAYRDLTFRESNASHVNFORPLLRKRVGGRMSAG 658
 788 RDESDIDISEASSVCSERSFSSSYTR-GNKSNY-----SLG-----GSHTRLDWSTQAP 836
 659 R-----RRSFDSQIQIGDISNFWGVPASLINALDGLNSSSDWCARVAA 703
 837 FDDITIIIOFCASHTMSERKQGLISLQVLADGKELTQOQKLCVLDMPKXMTDHTHY 896
 704 ENFLOTLLQ-----QG-----PKGAQEVITQSEFKYMKLFLRHLDDBHHKVA 744
 897 SLFLDTVTELIVANETSRRNGSSC-----LTRLFNKLGTDILNLSMHSKIWT 945
 745 QAALSTLADLI-----PSCRKPFESYMERVLPHVSRLL-IDPKEVVRPQCSST 791
 946 LQVHEHYPTQQLKELFRRIISDTQPTTTKTRIALILRLDILANTYCKSSDFPSDQOA 1005

DB 792 LEIVSKTYSDSLPALLRSL-DEQRP--VAKLAIVEFAINSFNRVAGNEPI-SGNSGI 847
 QY 1006 CERVLTKLAQADQKMELETSQARSCIVALYN-LNTEPQMTLLADLPKYVQDSARSCIH 1064
 DB 848 LKMLAKITPLTRD-KYTKLKEASITCISVYNNHYDAGLNLNLSLVEEQNSLRRLK 906
 QY 1065 -----SHNRROSQSCNSGANSPPSSPLSSSPKPLQSPS-----VGPRF-SLQ 1106
 DB 907 QYTRIEVDLNNYQSKKEQRIKSYDPSDA-IGTSSSEGYAGASKXNIFLGRYSGSID 965
 QY 1107 SHHQLSTSTSPRSRQSVQEGELFSELDIQNNIK-----TSEIRRCFGQY 1157
 DB 1158 QTAL---APNGFNGLQYHQQGQDSCASLSNSKTSQANTTOSNTPESATMELDNLER 1214
 1026 QNSISRTSPNOSSENILDD---DLSPPHLEKQNLNITSVDSLGRHENEVSRELDGHY 1082
 QY 1215 ERTTQNAKSPTDKAVITVSINMAENGELILIASLMESEVVRVALTLYKQPVELLQTS 1274
 DB 1083 MLTGIKNTTPESGPSIPQILHMINGSDGPS--KSGIQQLIEASV 1129
 QY 1275 TNLGICIKGNCGLPNKGFPSIMMLNILEAETDVVIAGLHYLSKIMRSNK--MRINW 1332
 DB 1130 AN-----EESWTYTFNQLTVLEVLDEDEFSIKELALSLISEMLKQKXAMEDS- 1180
 QY 1333 MHFELILIKTIQCYOH-----SKALRDIDSMIPRIAPSLPLDSINIVAPVIATGEP 1387
 DB 1181 ---VEIYEKLLHYSKQTPRVNSTEAEQCLTVVSQYDF---RCLSVIPLLVTEBEK 1233
 QY 1388 TNLCAIKILLEVTEHSGSEITDAHLDIVPNNARSADDTQSMVRAVFCIVLYFVLGE 1447
 DB 1234 TLVACINCLTFLVGRLSQEBLMDQSSFLPAVFAFGSQSADYKTVVFCIVDIYIMLGK 1293
 QY 1448 EKVPKSLVNPSTKRLNLY---IEKQRI 1474
 DB 1294 AFL-PYLEGNSTQVRLVTVIANRISQARN 1322

RESULT 13
 ADB63929
 ID ADB63929 standard; protein: 454 AA.
 XX
 AC ADB63929;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Human protein encoded by clone BRACE20079370.
 XX
 KW Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
 KW cell regeneration; membrane protein; signal transduction-related protein;
 KW transcription-related protein; osteoporosis; neurological disease;
 KW cancer; tumour.
 KW
 OS Homo sapiens.
 XX
 PN EP1308459-A2.
 XX
 PD 07-MAY-2003.
 XX
 XX 28-MAR-2002; 2002EP-00007401.
 XX
 XX 05-NOV-2001; 2001JP-00379298.
 PR 25-JAN-2002; 2002US-00350978.
 XX
 XX (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Isegaki T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
 PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuno Y,

DR WPI: 2003-450961/43.
 DR N-PSDB; ADB61959.
 XX New polynucleotides and polypeptides, useful for developing a diagnostic
 PT marker or medicines for regulation of their expression and activity, or
 PT as targets of gene therapy.
 PS Claim 1; Page; 222pp; English.
 XX The invention discloses a polynucleotide comprising a sequence selected
 CC from 1970 fully defined nucleotide sequences which encode novel
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
 CC or its partial peptide, an antibody binding to the polypeptide or peptide
 CC of the polynucleotide, immunologically assaying the polypeptide or
 CC peptide of the polynucleotide by contacting the polypeptide or peptide
 CC with the antibody of the encoded protein, and observing the binding
 CC between the two, a transformant carrying the polynucleotide in an
 CC expressible manner and an antisense polynucleotide. The oligonucleotide
 CC is useful as a primer for synthesizing the polynucleotide, or as a probe
 CC for detecting the polynucleotide. The polynucleotides and encoded
 CC proteins are useful as pharmaceutical agents and many disease-related
 CC genes may be included in them, for developing a diagnostic marker or
 CC medicines for regulation of their expression and activity, or as targets
 CC of gene therapy. The genes are involved in tissue and/or cell
 CC regeneration. Membrane proteins, signal transduction-related proteins,
 CC transcription-related proteins, disease-related proteins and genes
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
 CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
 CC the activity or expression of the encoded protein to treat diseases. The
 CC sequence presented is a protein of the invention. Note: Some of the
 CC sequence data for this patent is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.
 XX Sequence 454 AA:
 SQ
 Query Match 6.6%; Score 502.5; DB 7; Length 454;
 Best Local Similarity 32.5%; Pred. No. 1e-28;
 Matches 138; Conservative 77; Mismatches 156; Indels 54; Gaps 12;
 QY 261 PTRNKRPLHSAVSSSLRKPNDVDTGAGAVTMESSESEFEVVPOLNTHADMDIY 320
 DB 44 PANSARKP-----GSAGPKVGASKEGAGAVDEDFIKFTVPSIQIYSRELEBTL 98
 QY 321 KOVLITISDKNADMEKRDALKKIRALLI---LSYHTQPOFVAVOLKELISFVDILKEE 377
 DB 99 NKIREIISDDGHDQDQANALKKIRSLVAGAAGQDCCFQHLRLIDGALKLS-----AKD 153
 QY 378 LRSQVIREACITTIYAMSKTLNKLDAFCMSILHNLINLIONSAKVITASASTIAKYLKY 437
 DB 154 LRSQVIREACITTIYAMSKTLNKLDAFCMSILHNLINLIONSAKVITASASTIAKYLKY 213
 QY 438 THAPRLKIYDTLNOSSKSDIRSTLCELMVLLFEEMOTKALERNATVLRITLKKSIGDA 497
 DB 214 THVRLPLILINSNCT-SKSVFVRRSEFEFLDLLLOEMQTHSLERHAALVETIKKGIDA 272
 QY 498 DCDARRHRVAYMAFRHFPPELLADQIYGTLDIAAORALEREEREGGGGTGTGTAPEY 557
 DB 273 DAARVREARKTYMGLRNHFPGEAETLYNSLEPSYOKSLQTYLKSSG-----SVASLPDS 326
 QY 558 RRTYSRIGRTGTLQKTPPSMRSSISAVDTAAORAKVRAOYTLVSRORRKLPGPNNSQAS 617
 DB 327 DRSSS---SSGESLNRPFSSSKMSTANPSTVA---GRVSA-----GSSKAS 365
 QY 618 MTGAAGSSLPFRP---LNSNSGGTPTATTPGVSVPFRPRGAGVSSQSPGSRSTSPITKL 673
 DB 366 ---SLPESLORSREDIDVNAAGAKAHAAAGQSV---RSGRIGAGALNAGSYA---SLGK 416
 QY 674 RDQYG 678
 DB 417 RTDYG 421

RESULT 14
 ABB81781
 ID ABB81781 standard; protein; 432 AA.
 XX
 XX ABB81781;
 AC
 XX
 DT 24-SEP-2002 (first entry)
 DE
 XX Proteasome subunit nu 47.52.
 DE
 XX Proteasome subunit nu 47.52; protein metabolic disorder.
 KM
 XX Unidentified.
 OS
 XX CN139447-A.
 PN
 XX 13-MAR-2002.
 PD
 XX 21-AUG-2000; 2000CN-00119655.
 PF
 XX 21-AUG-2000; 2000CN-00119655.
 PR
 XX 21-AUG-2000; 2000CN-00119655.
 XX
 PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
 XX
 PI Mao Y, Xie Y;
 PT
 DR WPI: 2002-455349/49.
 DR N-PSDB; ABB83145.
 XX
 PT New polypeptide-proteasome subunit nu 47.52 and polynucleotide for
 PT encoding such polypeptide.
 PS
 PS Claim 1; Page 26-27 (Disclosure); 33pp; Chinese.
 CC The sequence represents the novel proteasome subunit nu 47.52 of the
 CC invention. The invention relates to the novel proteasome subunit nu
 CC 47.52, and the polynucleotides encoding it. The present invention also
 CC discloses the method of applying the polypeptide in treating various
 CC diseases, such as protein metabolic disorder. The antagonist of the
 CC polypeptide and its treatment effect are also disclosed
 CC
 XX Sequence 432 AA:
 SQ
 Query Match 6.5%; Score 497.5; DB 5; Length 432;
 Best Local Similarity 32.7%; Pred. No. 2.3e-28;
 Matches 136; Conservative 75; Mismatches 154; Indels 51; Gaps 11;
 QY 237 SALK---NINNGVGLDEADNIGLERPPTRMIRPLHSAVSSSLRKPNNV-DVTGDAG 291
 DB 31 SAEKVPAPKTSN-----PANSARKP-----GSAGPKVGASKEGAG 70
 QY 292 AYTMESSESEFEVVPOLNTHADMDIYKOVLITISDKNADMEKRDALKKIRALLI-- 349
 DB 71 TVBEDDPIKFTVPSIQIYSRELEBTLNKIREIISDDGHDQDQANALKKIRSLVAG 130
 QY 350 -LSYHTQPOFVAVOLKELISFVDILKEELRSQVIREACITTIYAMSKTLNKLDAFCMSI 408
 DB 131 AAGYDCFFQHLRLIDGALKLS-----AKDLRSQVIREACITTIYAMSKTLNKLDAFCMSI 185
 QY 409 LEHNLINLIONSASVITASASTIAKYLKYTHAPRLKIYDTLNOSSKSDIRSTLCELMV 468
 DB 186 VPTLFLNVPNSAKVITASASTIAKYLKYTHAPRLKIYDTLNOSSKSDIRSTLCELMV 244
 QY 469 LTFEEMOTKALERNATVLRITLKKSIGDADCDARRHRVAYMAFRHFPPELLADQIYGTLD 528
 DB 245 LILQEMQTHSLERHAALVETIKKGIDAAARVREARKTYMGLRNHFPGEAETLYNSLE 304
 QY 529 IAAQRALEREREEREGGGGTGTGTAPEYTRTYSRIGRTGTLQKTPPSMRSSISAVDTAA 588
 DB 305 PSYOKSLQTYLKSSG-----SVASLPDSRSSS---SSGESLNRPFSSSKMSTANPSTVA 355
 QY 589 ---AORAKVRAOYTLVSRORRKLPGPNNSQASMTGAAA---SGSLPRPLNNS 636

Db 356 GRVSGSSKASLFGSLRRSRSDIGVNAAGAKAHHAAGOSVRSGLGALNNGS 411

RESULT 15
ADJ68457
ID ADJ68457 standard; protein; 1972 AA.

XX AC ADJ68457;
XX DT 06-MAY-2004 (first entry)
XX DE Human heart mitochondrial protein as a therapeutic target SeqID263.
XX KM mitochondrial; human; screening assay; diabetes mellitus;
XX KM Huntington's disease; osteoarthritis;
XX KM Leber's hereditary optic neuropathy; LHON;
XX KM mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
XX KM myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
XX KM neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarrhythmic;
XX KM osteopathic; ophthalmological; cytoskeletal.
XX OS Homo sapiens.
XX PN NO2003087768-A2.
XX PD 23-OCT-2003;
XX PF 04-APR-2003; 2003MO-US010870.
XX PR 12-APR-2002; 2002US-0372843P.
XX PR 17-JUN-2002; 2002US-038987P.
XX PR 20-SEP-2002; 2002US-0412418P.
XX PA (MITO-) BUCK INST AGE RES.
XX PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
XX PI Warnock DE;
XX DR WPI; 2003-845369/78.
XX PT Identifying a mitochondrial target for drug screening assays and for
XX PT treating diseases associated with altered mitochondrial function,
XX PT comprises detecting a modified polypeptide in a sample and correlating
XX PT with the disease.
XX PS Claim 1; SEQ ID NO 263; 180pp; English.
XX CC This invention relates to novel mitochondrial targets that can be used
XX CC for therapeutic intervention in treating a disease associated with
XX CC altered mitochondrial function. Specifically, it refers to a method for
XX CC identifying proteins of the human heart mitochondrial proteome that are
XX CC useful for drug screening assays, as well as therapeutic targets. The
XX CC present invention describes a method for identifying such proteins that
XX CC can be used in the treatment of various diseases associated with altered
XX CC mitochondrial function including diabetes mellitus, Huntington's disease,
XX CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
XX CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
XX CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
XX CC compositions have neuroprotective, nootropic, antidiabetic, and
XX CC anticonvulsant, antiarrhythmic, osteopathic, ophthalmological and
XX CC cytoskeletal activities. This polypeptide sequence is a human heart
XX CC mitochondrial protein of the invention.
XX SO Sequence 1972 AA;

Query Match 4.3%; Score 328; DB 7; Length 1972;
Best Local Similarity 18.7%; Pred. No. 1,7e-14;
Matches 278; Conservative 253; Mismatches 543; Indels 416; Gaps 68;

QY 61 SHFPIAKSLSEFELIKRLGSDPNATATVLPVIVIRLSDRDTREKAQLLRDMEH 120
DB 653 TNGVWQMKLHYA-LIAQKGNFSKTSIAQVVLDDGLVVKIDVK--GNNAKEMTAIAEA 709

QY 121 RVLPQALIDKLATSCFKRKNAAVEEFLQTI VNALHEXYTQQLSVRYVTPVYCALLGDP 180
DB 710 CMLPMTA--EGVSNMAFSQKNPKNOSETLNLWLSNAIKERFGGLNVKAPISNVKTAALAT 767
QY 181 TNNVREAAIQTVLELYKKVGDRLRPDLRM--DDVPASLAMLKQKPDQVKGGLLPSA 238
DB 768 NPAYRTAATLLTGVMVLYG---PSLRPFEDKPA-LLSQIDAEFB--KQGGSPAP 820
QY 239 LKNTNGNV-GLDEADNIGLRERPTRMIKRPLSAVSSSLRPKNVDYTGAGAVTMS 297
DB 821 TRGISKHSSTGDBED-----GDEPDGSDNDV----- 849
QY 298 FESSFEVPEQLNIFPAKMDIYKQVLYTISPKADMEKRVDAKKIRALLSYHTQO 357
DB 850 ----DLPRTEI-----SDKITSELYSGIKGN--WKIRKGLDEVAQII-----NDAK 892
QY 358 FVAVQLKELSLFVDILKEELRSQVIRACITIAVMSKTLRKLDAPCWSILEHJINLQ 417
DB 893 FIQPNIGELPTA---LGRNLDS-----NKI-----LVQQTINLQ 925
QY 418 NSAKVIASASTIALKYIIKYTHAPLLKLYTDTLNSKSDIRSTLCMLVLLFEMQK 477
DB 926 Q---LAVMGPNIKQHVNLQIP-IITVLGDSKNVRAALATV----- 965
QY 478 ALERNATVLRDTLKKSIGDADCDARRHSRYAVAFRRHPELADQIYGTLDIAQRALE 537
DB 966 ---NMAEQTMKEWLEGEEDLSE-----LKENPFILROELGWL---AEK--- 1005
QY 538 EREGGGGGTGCTGTAPETRTVSRIGRTPTLQKPPMSMSIAVDPAANOR---AKV 594
DB 1006 -----LPLRST-----PDLILCVPHLLSCLDEKRDVAK 1036
QY 595 RAQYTL-----YSROKPLG---PNNSQASMTGAASGLSPRLNSGGTPART 643
DB 1037 KAQDALPFFMHLGTEKAKATGKXKLPISKO-----VLAMEKAKYNNAPAKAP 1086
QY 644 PGSVTPRPRGAGVSQSGPGRST---SPSTKLDQYGIGNYRGATGALPKASGIP 699
DB 1087 PTKATSKPWGSGAPAKFPASAPADCISSSTPEKPD-----PKKA----- 1127
QY 700 RSTASRETSPTRSSGGLMKRMYSTGASRRTPRNNVPRSPARLLAQSREAHTLG 759
DB 1128 -----KAPGASSKAKSAOG--KMPKSTK-----LKEDE----- 1154
QY 760 VGDDGQPDYV---SGDYMSGGMRGKLM-----GRDE-SDDIDSEASVCS---ER 805
DB 1155 --DKSGPIFIVPNKQKQRMKQEK-GLKVLKKNFTTPREVIHQKTMSSCVAAKVLQDE 1211
QY 806 SFDSSYTRGNKSNVLSGSHRTLDMWSTORAPPDIIETIIQFCASTHWSERKGLISLQY 865
DB 1212 MHSDFOHNKKA-----LAVWD-----HLESEKGVIGC--- 1241
QY 866 LADGKELTQQLKCVLDMFRKMPMDTHPKVSLPDTVELLYVANE---TSRNGSSSC 922
DB 1242 -----DDLTKMLTLRFPTNLSVLKMALEYKLTLTSEEYHHTENASSF 1290
QY 923 LTRLENKLG--VDLNSMHSKIMKTLOVNEHEFTQLOKELFRJISDSTPTTKTRIA 980
DB 1291 IYLVVVKGERPDVIRKQVRAILNRMCLVY-----PASKKPFIMESTKSKSQRAE 1343
QY 981 IIRFLTDLANTYCKSSDPSPDOSQACERT---VLKLAQLADQKSMELRSQARSCVALLY 1037
DB 1344 CLEELGLAVESY-----GMNVCOPFGKALKEIVHIGDRNVAVNAALNTIVTY 1394
QY 1038 NLNTPQMTLLADLPKVQDSARSCTHSHMRQSGSCNSGANSPPSSPPLSSPPLQSP 1097
DB 1395 NVHGQVFKLIENL---SEKMSMLBERIK-----SARPSAAPTQVKEKFORAO 1443
QY 1098 SV-----GPFASLQSHHQLSISSTSPRSQSVYQELLFSSSELDIQNIQKT--- 1145
DB 1444 NISSNANMLRKRPADWMSKLNQARSMSGHPEAAQ-VWRRE--FOLDDEIENDQTYVC 1500

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: November 29, 2005, 14:27:58 ; Search time 12424 Seconds
(without alignments)
19375.365 Million cell updates/sec

Title: US-10-030-850-1
Perfect score: 5145
Sequence: 1 gaattcgagcagcgcgc.....ggattaaatcaataaaaa 5145

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb-est1:*
2: gb-est2:*
3: gb-est3:*
4: gb-est4:*
5: gb-est5:*
6: gb-est6:*
7: gb-est7:*
8: gb-est8:*
9: gb-est9:*
10: gb-est10:*
11: gb-est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	925.2	18.0	931	1	A1513688 GH26741.5
2	836.6	16.3	932	2	BG633422 GH26741.3
3	803.2	15.6	843	10	CNS000AD8
4	762.8	14.8	766	1	A1405248 GH25189.5
5	741.2	14.4	856	2	BG633353 GH25189.3
6	719	14.0	720	2	BF485581 AT19621.5
7	716	13.9	716	7	CK658052 LP14575.5
8	692	13.4	692	2	BF504405 AT05759.5
9	645	12.5	656	1	A1297625 LP11963.5
10	619.6	12.0	642	10	AG977859 Drosophila
11	604.4	11.7	606	2	BG634723 AT31034.5
12	597	11.6	626	2	BF493211 AT01524.5
13	586	11.5	615	2	BF491691 AT28490.5
14	581	11.4	631	2	BF491691 AT28490.5
15	587.2	11.4	631	2	BF491691 AT28490.5
16	580	11.3	580	2	BF491691 AT28490.5
17	574.2	11.2	579	2	BF502300 AT17981.5
18	552	10.7	576	10	AG951156 Drosophila
19	551	10.7	551	1	A1543213 Drosophila
20	543	10.6	566	1	A1294703 LP08134.5
21	530.6	10.3	559	1	A1945450 Drosophila
22	517.8	10.1	523	1	A1945833 Drosophila

23	509.4	9.9	511	7	CO314038
24	507.6	9.9	561	10	AG973742
25	500	9.7	500	1	A1456404
26	486.6	9.5	629	10	AG912861
27	476.2	9.3	512	2	BF492167
28	473.4	9.2	475	7	CO307834
29	471	9.2	494	7	CO305613
30	469	9.1	518	2	BF486127
31	468	8.9	484	10	AG962276
32	450	8.7	450	7	CO284242
33	445.6	8.7	861	10	CNS01100
34	435	8.5	498	2	BG639454
35	425.4	8.3	427	7	CO289204
36	420	8.2	1101	10	CNS014X8
37	399.2	7.8	437	1	A1946614
38	394.8	7.7	429	7	CO315184
39	382.6	7.4	412	7	CO270671
40	379	7.4	572	1	AA541072
41	360	7.0	403	1	AM943097
42	359	7.0	359	7	CO263714
43	298.2	5.8	303	7	CO286919
44	279.2	5.4	285	7	CO268707
45	274	5.3	460	2	BE977670

ALIGNMENTS

RESULT 1
LOCUS A1513688
DEFINITION GH26741.5prime GH Drosophila melanogaster head pot2 Drosophila
melanogaster cDNA clone GH26741.5, mRNA sequence.
A1513688
ACCESSION A1513688
VERSION A1513688.2 GI:13759214
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 911)
REFERENCE Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G., Lewis, S. and Rubin, G.M.
BDGP/HMI Drosophila EST Project
Unpublished (2001)
On Mar 16, 1999 this sequence version replaced gi:4417834.
Other ESTs: GH26741.3prime
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
hit genomic AE005593: arm:3L (2079555,2112669)
estimated-cyto:78A2-78C7: 04/10/2001
Plate: GH.267 row: D column: 5
High quality sequence stop: 616
POLYA=No.

FEATURES

source

Location/Qualifiers
1..931
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="GH26741"
/sex="male and female"
/dev stage="adult"
/lab_host="DHS - alpha"
/clone_lib="GH Drosophila melanogaster head pot2"
/note="Organ: head; Vector: POT2; Site: 1; EcoRI; Site 2; XhoI; Sized fractionated cDNAs were directly ligated into
POT2. Plasmid cDNA library."
ORIGIN


```

RESULT 4
AL05248      766 bp      mRNA      linear      EST 23-APR-2001
LOCUS       GH25189.5prime GH Drosophila melanogaster head pot2 Drosophila
DEFINITION  melanogaster cDNA clone GH25189 5, mRNA sequence.
ACCESSION   AL05248
VERSION     AL05248.2  GI:13759105
KEYWORDS
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM    Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 766)
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
Lewis, S. and Rubin, G. M.
BDGP/HMT Drosophila EST Project
Unpublished (2001)
On Feb 8, 1999 this sequence version replaced gi:4248335.
Other ESTs: GH25189.3prime
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AB003593: arm:3L [20799555,21112669]
estimated-cyto:78A2-78C7: 04/10/2001
Plate: GH.251 row: H column: 5
High quality sequence stop: 634
POLYA=No.

FEATURES
    source
        1..766
            /location/Qualifiers
            /organism="Drosophila melanogaster"
            /mol_type="mRNA"
            /db_xref="taxon:7227"
            /clone="GH25189"
            /sex="male and female"
            /dev_stage="adult"
            /lab_host="DHS - alpha"
            /clone_1lb="GH Drosophila melanogaster head pot2"
            /note="Organ: head; Vector: pot2; Site_1: EcoRI; Site_2:
            XhoI; Sized fractionated cDNAs were directly ligated into
            pot2. Plasmid cDNA library."

ORIGIN
Query Match      14.8%; Score 762.8; DB 1; Length 766;
Best Local Similarity 99.7%; Pred. No. 5.1e-212;
Matches 764; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1257  AGAATTAAGTATGATGCTTCTGCTGAGACATTTTGAGACACCTGATTATTATACG 1316
      |||||
      1  AGAATTAAGTATGATGCTTCTGCTGAGACATTTTGAGACACCTGATTATTATACG 60

1317  AACAGCGGAGAGTATGATGCTTCTGCTGAGACATTTTGAGACACCTGATTATTATACG 1376
      |||||
      61  AACAGCGGAGAGTATGATGCTTCTGCTGAGACATTTTGAGACACCTGATTATTATACG 120

1377  ACACATGACCAAGCTGCTCAAGATCTACAGACACTCTGAATCAATCAAGTCGAG 1436
      |||||
      121  ACACATGACCAAGCTGCTCAAGATCTACAGACACTCTGAATCAATCAAGTCGAG 180

1437  GACATTAAGTATGATGCTTCTGCTGAGACATTTTGAGACACCTGATTATTATACG 1496
      |||||
      181  GACATTAAGTATGATGCTTCTGCTGAGACATTTTGAGACACCTGATTATTATACG 240

1497  GCGTTGAAAGGAATGCCACCGTACTAAGGACACCTTAAAAAATCCATTGGCGATGCA 1556
      |||||
      241  GCGTTGAAAGGAATGCCACCGTACTAAGGACACCTTAAAAAATCCATTGGCGATGCA 300

1557  GACATGAGTACGACCGCGCATTCAGATACGGCTATTGGGCTTCAGAGGCGATCTTCCA 1616
      |||||
      301  GACATGAGTACGACCGCGCATTCAGATACGGCTATTGGGCTTCAGAGGCGATCTTCCA 360
  
```

```

1617  GAGCTGGGGATCAAAATATATAGAACATTAGACATAGCTGCCACCGCGATTAGAAAG 1676
      |||||
      361  GAGCTGGGGATCAAAATATATAGAACATTAGACATAGCTGCCACCGCGATTAGAAAG 420

1677  GAACGAGGGGGCGGAGAGAGAGAACTGTACTGGGACTGGGACTGCACTGAAACG 1736
      |||||
      421  GAACGAGGGGGCGGAGAGAGAGAACTGTACTGGGACTGGGACTGCACTGAAACG 480

1737  AGAGCACTGATCCCGCATTTGGCGGAAACACCTGGAACCTTGCAAAAGCCAGCGCTAGT 1796
      |||||
      481  AGAGCACTGATCCCGCATTTGGCGGAAACACCTGGAACCTTGCAAAAGCCAGCGCTAGT 540

1797  ATGAGATCATTTTACGCGGTGACACTGCGGCTCTCAACGAGCCAAAGTTAGACGCA 1856
      |||||
      541  ATGAGATCATTTTACGCGGTGACACTGCGGCTCTCAACGAGCCAAAGTTAGACGCA 600

1857  TATACACTATATTCAGGCAAAAGAAACCTTTAGACCTATATTTCCACAGCATCG 1916
      |||||
      601  TATACACTATATTCAGGCAAAAGAAACCTTTAGACCTATATTTCCACAGCATCG 660

1917  ATGACAGAGAGAGAGAGAGATGATCACTGCGCCAGACCTCGCTGAATTCATAGCGGT 1976
      |||||
      661  ATGACAGAGAGAGAGAGAGATGATCACTGCGCCAGACCTCGCTGAATTCATAGCGGT 720

1977  GGACACCAAGCTACACGCGCGGATCGTTACACACGCGCCCGCG 2022
      |||||
      721  GGACACCAAGCTACACGCGCGGATCGTTACACACGCGCCCGCG 766

RESULT 5
B6633353/C      856 bp      mRNA      linear      EST 23-APR-2001
LOCUS       GH25189.3prime GH Drosophila melanogaster head pot2 Drosophila
DEFINITION  melanogaster cDNA clone GH25189 3, mRNA sequence.
ACCESSION   B6633353
VERSION     B6633353.1  GI:13759104
KEYWORDS
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM    Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 856)
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
Lewis, S. and Rubin, G. M.
BDGP/HMT Drosophila EST Project
Unpublished (2001)
Other ESTs: GH25189.5prime
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or
more T residues at the beginning of the sequence, this clone was
polyadenylated. The resulting Poly-T sequence has been removed. hit
genomic AB003593: arm:3L [20799555,21112669]
estimated-cyto:78A2-78C7: 04/10/2001
Plate: GH.251 row: H column: 5
High quality sequence stop: 363.

FEATURES
    source
        1..856
            /location/Qualifiers
            /organism="Drosophila melanogaster"
            /mol_type="mRNA"
            /db_xref="taxon:7227"
            /clone="GH25189"
            /sex="male and female"
            /dev_stage="adult"
            /lab_host="DHS - alpha"
            /clone_1lb="GH Drosophila melanogaster head pot2"
            /note="Organ: head; Vector: pot2; Site_1: EcoRI; Site_2:
            XhoI; Sized fractionated cDNAs were directly ligated into
            pot2. Plasmid cDNA library."
  
```

XhoI; sized fractionated cDNAs were directly ligated into
pOT2. Plasmid cDNA library."

Query Match	14.4%	Score 741.2	DB 2	Length 856
Best Local Similarity	98.7%	Pred. No. 1.2e-205		
Matches 758	Conservative	0	Mismatches 8	Indels 2
			Gaps	1
Qy	4380	GTCAAGCTGTACTTGTCTTGGGGGAGAGAAAGTCAAGCCGAAGCTGTAGTCTAAAT	4439	
Db	856	GTAAGAGCTGTACTGTGTCTTGGGGGAGAGAAAGTCAAGCCGAAGCTGTAGTCTAAAT	797	
Qy	4440	CCCAAGCAAGTTAAGGCTCTCAACGCTGATCGAAGCAAGCAAGCAAGTCAAGTCAAGTGG	4499	
Db	796	CCCAAGCAAGTTAAGGCTCTCTCAAGCTGTACATCGAAGCAAGCAAGCAAGTCAAGTGG	737	
Qy	4500	GGAGAGAGCTCTACAAAGAACTCTCCGCGGCAATGTCGATCATATTGGGAGCCCTTA	4559	
Db	736	GGAGAGAGCTCTACCAAGAACTCTCCGCGGCAATGTCGATCATATTGGGAGCCCTTA	677	
Qy	4560	ATAGAGTTCCTGC - TCGTGCACCAACAAACAGACAGAGCGGGTGTCTCCCTGGG	4617	
Db	676	ATAGAGTTCCTGCATCAGACACACACAAACAGACAGAGCGGGTGTCTCCCTGGG	617	
Qy	4618	CTGAAGAGAGTGAAGAGCGCGGCACTTAATTAATATTATTAATTAATTAATTAATTAAT	4677	
Db	616	CTGAAGAGAGTGAAGAGCGCGGCACTTAATTAATTAATTAATTAATTAATTAATTAAT	557	
Qy	4678	TTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	4737	
Db	556	TTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	497	
Qy	4738	CAGATGCAATGCGGCGCCGAAAGAAAGTCAAGGCGCATCGCCATCTAATGAGACAGA	4797	
Db	496	CAGATGCAATGCGGCGCCGAAAGAAAGTCAAGGCGCATCGCCATCTAATGAGACAGA	437	
Qy	4798	AAACTATTATTAATCATTAACGGGAACTAAGCGAGTAATTCGCAAGTGTACTAAATTTG	4857	
Db	436	AAACTATTATTAATCATTAACGGGAACTAAGCGAGTAATTCGCAAGTGTACTAAATTTG	377	
Qy	4858	TAGCCAACTCCGCAATCTCTACTTCTCACATCATGTTGACGTTCTTTTACCAGCTCCCTTA	4917	
Db	376	TAGCCAACTCCGCAATCTCTACTTCTCACATCATGTTGACGTTCTTTTACCAGCTCCCTTA	317	
Qy	4918	CTATTAAGCGCGAATCTGTTTAATTTGTAAGCGCTATATACGCGCTTTATGGAATCTAG	4977	
Db	316	CTATTAAGCGCGAATCTGTTTAATTTGTAAGCGCTATATATACGCGCTTTATGGAATCTAG	257	
Qy	4978	ACGTGTGCTTTCTACCTTTGTTGAGGCGACTTTGTAACCTTAATTTGTTCCGCTCA	5037	
Db	256	ACGTGTGCTTTCTACCTTTGTTGAGGCGACTTTGTAACCTTAATTTGTTCCGCTCA	197	
Qy	5038	CTAAATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	5097	
Db	196	CTAAATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	137	
Qy	5098	TTAATTTAAATACACAAAGCCACACAAAGAGTTAAATCAATTAATAAA 5145		
Db	136	TTAATTTAAATACACAAAGCCACACAAAGAGTTAAATCAATTAATAAA 89		
RESULT 6				
BF485581				
LOCUS	BF485581	720 bp	mRNA	linear
DEFINITION	AT19621.5prime AT Drosophila melanogaster adult testes p0H87			
ACCESSION	BF485581			
VERSION	BF485581.2	GI:13697518		
KEYWORDS	EST.			
SOURCE	Drosophila melanogaster (fruit fly)			
ORGANISM	Drosophila melanogaster			
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			

REFERENCE
AUTHORS

Ephydroidea; Drosophilidae; Drosophila
1 (bases 1 to 720)
Stapleton, M., Brokstein, P., Hong, L., Ac-
berman, B., Carlson, J., Champe, M., Chavez,

TITLE BDGP/HHMI AT Drosophila EST Project
JOURNAL Unpublished (2000)
COMMENT On Dec 6, 2000 this sequence version replaced gi:11568918
Contact: Stanleron, M

Lawrence Berkeley National Lab
One Cyclotron Rd.
Berkeley, CA 94720, USA
Fax: 510 486 6796
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AB003593: arm:3L [20799555,21112663]
estimated-cyco:78A2:78C7: 04/09/2001
Plate: AT.196 row: B column: 9
High quality sequence stop: 647.
Identification: 040115f9e

ORIGIN

Query	Match	Similarity	14.0%	Score	719	DB	21	Length	720
Best Local	Similarity	99.9%	Pred.	No.	319	159			
Matches	719	Conservative	0	Mismatches	1	Indels	0	Gaps	0
Qy	3393	CACCAACTTGCATACGCTCTACTAGTCCAGCGCTCCCGGACGATTCCTCGTGGACGACGAG	3452						
Dp	1	CACCAACTTGCATACGCTCTACTAGTCCAGCGCTCCCGGACGATTCCTCGTGGACGAGG	60						
Qy	3453	CTGCTCTTTTCTCGGAGCTGGACATTCAGACCAACATTCAGAAAGCGTCGAGAGATC	3512						
Dp	61	CTGCTCTTTTCTCGGAGCTGGACATTCAGACCAACATTCAGAAAGCGTCGAGAGATC	120						
Qy	3513	CGGCACTGCTTGGCGCGTCACTATCCAGACGCGCTGGCGCCCAATGGCTTCAATGGANAC	3572						
Dp	121	CGGCACTGCTTGGCGCGTCACTATCCAGACGCGCTGGCGCCCAATGGCTTCAATGGACAC	180						
Qy	3573	TTGACAGTATCAGATCAGGCGCCAAACAGATTGCTGTGATCCCTGCTTCCAACTCCAAG	3633						
Dp	181	TTGACAGTATCAGATCAGGCGCCAAACAGATTGCTGTGATCCCTGCTTCCAACTCCAAG	240						
Qy	3633	ACGCATTCGTGGCGCAACACTACCGAGTCAATATACCTGATCAGCAACATGAGGCTG	3693						
Dp	241	ACGCATTCGTGGCGCAACACTACCGAGTCAATATACCTGATCAGCAACATGAGGCTG	300						
Qy	3693	GATATCTGGAGCGGAAAGGACCACTAGAAACGCAAGTCAACCACTGACGATGCCAAG	3753						
Dp	301	GATATCTGGAGCGGAAAGGACCACTAGAAACGCAAGTCAACCACTGACGATGCCAAG	360						
Qy	3753	GTGATCAGCGTCTCGATTAATATGAGCTGAAAATGAGAGCTGATATCTGGCCAGCAACTG	3813						
Dp	361	GTGATCAGCGTCTCGATTAATATGAGCTGAAAATGAGAGCTGATATCTGGCCAGCAACTG	420						
Qy	3813	ATGGAGACGGAAGTGTCGCTGAGGCTTACGCTAACAAAGATCAAGCCGTCGAGTTG	3872						

Db	421	ATGGAGAGGGAAGTGTGTGTGTGGCTTGACGCTTAACAAGAATCACCCGTGCAGTTG	480
Qy	3873	CTTCAGACGTCCTACTACTAACCTGGGGATTGTCATCAAGGGCGAAATGTGTAGCTGCC	3932
Db	481	CTTCAGACGTCCTACTACTAACCTGGGGATTGTCATCAAGGGCGGAAACTGTAGAGTGCC	540
Qy	3933	AATAAGCATTNAGTCCATCATCGCAATGCTGCTTAACTTTCTGAGACGGAGCATACG	3992
Db	541	AATAAGCACTTTAGTTCGATCATCGCAATGCTGCTTAACTTTCTGAGACGGAGCATACG	600
Qy	3993	GACGGTGATCGCTGGCGCTGACGTCGTCAGTAGATTATGAGAGCAACAAATGGCT	4052
Db	601	GACGGTGATCGCTGGCGCTGACGTCGTCAGTAGATTATGAGAGCAACAAATGGCT	660
Qy	4053	CACAACTGGATGCACTTTCTAGAGCTGATTTTGTCTGAAGATCATCGATGCTATCAAC	4112
Db	661	CACAACTGGATGCACTTTCTAGAGCTGATTTTGTCTGAAGATCATCGATGCTATCAAC	720
RESULT 7			
LOCUS	CK658052	716 bp mRNA linear EST 30-JAN-2004	
DEFINITION	LPI4575:5prime LP Drosophila melanogaster larval-early pupal POT2		
ACCSSION	CK658052		
VERSION	CK658052.1 GI:41401577		
KEYWORDS	EST.		
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Drosophila melanogaster		
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
AUTHORS	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
JOURNAL	Ephydriidea; Drosophilidae; Drosophila.		
TITLE	1 (bases 1 to 716)		
COMMENT	Havrey,D., Broksrein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G., Lewis,S. and Rubin,G.M. BDGP/HMT Drosophila EST Project Unpublished (2001) Contact: Stapleton, M. BDGP Lawrence Berkeley National Lab One Cyclotron Rd, Berkeley, CA 94720, USA Fax: 510 486 6798 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu Plate: LP.145 row: G column: 3 High quality sequence stop: 562. Location/Qualifiers 1..716 /organism="Drosophila melanogaster" /mol_type="mRNA" /db_xref="taxon:7227" /clone="LPI4575" /sex="male and female" /dev_stage="larvae-pupae" /lab_host="DH5-alpha" /clone_lib="LP Drosophila melanogaster larval-early pupal POT2" /note="Organ: whole body; Vector: pOT2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into POT2. Plasmid cDNA library."		
FEATURES			
source			
ORIGIN			
Query Match	13.9%	Score 716;	DB 7; Length 716;
Matches Local Similarity	100.0%; Pred. No. 3e-198;		
Matches 716; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
Qy	881	CGTTTCGTCATCCTGCGCCCAAAACCAATGTGAACGATGTGACCGGTGATGCCGGCGC	940
Db	1	CGTTTCGTCATCCTGCGCCCAAAACCAATGTGAACGATGTGACCGGTGATGCCGGCGC	60
Qy	941	CGTAACTGGAATTTTTCGAATCTAGCTTTGAGGTGTCCCGCAATTGAACATCTTCCA	1000

[illegible]

estimated-cyto:78A2-78C7: 04/07/2001
 Plate: At.57 row: E column: 11
 High quality sequence stop: 652.
 location/Qualifiers

FEATURES

SOURCE

1. 692
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /clone="AT05759"
 /sex="male"
 /dev_stage="0-3 day old Ore-R males"
 /lab_host="Plates At.10-At.120: DHS-alpha. Plates
 At.121-At.319: DHS-alpha Tona"
 /clone_lib="At Drosophila melanogaster adult testes potB7"
 /note="Organ: ADULT testes; Vector: potB7; Site: 1: EcorI;
 Site: 2: XhoI; The mRNA for the testis library was made
 from testes and seminal vesicles hand dissected from 0-3
 day old Ore-R males. RNA kindly provided by the lab of
 Margaret Fuller. Sized fractionated cDNAs were directly
 ligated into potB7. Plasmid cDNA library."

ORIGIN

Query Match 13.4%; Score 692; DB 2; Length 692;
 Best Local Similarity 100.0%; Pred. No. 3.4e-191; Indels 0; Gaps 0;
 Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY CGGAGATCCACAGTTAATGTAGGAGGAGGCGGCGCATCCAAACGCTAGTGAATCTACAA 655
 DB 1 CGGAGATCCACAGTTAATGTAGGAGGAGGCGGCGCATCCAAACGCTAGTGAATCTACAA 60
 OY 656 GCATGTAGGAGATCGATTGCGCCAGACCTCGTGCATGACGATGTTCTCTCGCA 715
 DB 61 GCATGTAGGAGATCGATTGCGCCAGACCTCGTGCATGACGATGTTCTCTCGCA 120
 OY 716 ATTGCTATGTTGGAGCAAAAGTTGGACCAAGTCAACGAGGGCTTACTGCTACCTTC 775
 DB 121 ATTGCTATGTTGGAGCAAAAGTTGGACCAAGTCAACGAGGGCTTACTGCTACCTTC 180
 OY 776 AGCCCTTAAAAACAGATGGAATGGAATGAGGCTTGGACGAGCGGCAATATTGGGTT 835
 DB 181 AGCCCTTAAAAACAGATGGAATGGAATGAGGCTTGGACGAGCGGCAATATTGGGTT 240
 OY 836 GAGGAGCGACCCACAGAGATGTTAAGCGGCGCATCACTCGCGCGTTTCTCACT 895
 DB 241 GAGGAGCGACCCACAGAGATGTTAAGCGGCGCATCACTCGCGCGTTTCTCACT 300
 OY 896 GCGCCCAAAACCAATGTGAACGATGTACCGGTGATGCGGCGCGGTAAACATGGAATC 955
 DB 301 GCGCCCAAAACCAATGTGAACGATGTACCGGTGATGCGGCGCGGTAAACATGGAATC 360
 OY 956 TTTTCAATTTAGCTTGAAGTGTGCGGCAATTTGAACATCTTCAACGTAAGACATGGA 1015
 DB 361 TTTTCAATTTAGCTTGAAGTGTGCGGCAATTTGAACATCTTCAACGTAAGACATGGA 420
 OY 1016 CGATATCTCAAGCAAGTACTAGTATCACTAGTATTAATAAAGCAAGTGGAGAAAG 1075
 DB 421 CGATATCTCAAGCAAGTACTAGTATCACTAGTATTAATAAAGCAAGTGGAGAAAG 480
 OY 1076 TGTGATGCTTCTCAAGAAATCAGGCAATTTGCTCAATTTCACTCACTCAAGCCGA 1135
 DB 481 TGTGATGCTTCTCAAGAAATCAGGCAATTTGCTCAATTTCACTCACTCAAGCCGA 540
 OY 1136 GTTTGTCGCTGTACACTAAAGAAATTTGTTAAGCTTGTGGAATCTCTCAAGAGAGA 1195
 DB 541 GTTTGTCGCTGTACACTAAAGAAATTTGTTAAGCTTGTGGAATCTCTCAAGAGAGA 600
 OY 1196 ACTACATCAACAGTATCCGCGAGGCGTGCATCACCATGCTTACATGTCTAAGACGCT 1255
 DB 601 ACTACATCAACAGTATCCGCGAGGCGTGCATCACCATGCTTACATGTCTAAGACGCT 660
 OY 1256 GAGAAATTAACATGATGCTTGTGCTGGAGCA 1287
 DB 661 GAGAAATTAACATGATGCTTGTGCTGGAGCA 692

RESULT 9 656 bp mRNA linear EST 23-Apr-2001
 A1297625
 LOCUS
 DEFINITION
 Drosophila melanogaster CDNA clone Lp11963 5, mRNA sequence.

ACCESSION
 A1297625
 VERSION
 A1297625.1 GI:3947032
 KEYWORDS
 EST,
 Drosophila melanogaster (fruit fly)
 ORGANISM
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
 1 (bases 1 to 656)
 Harvey, D., Brooksstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
 Lewis, S., and Rubin, G.M.

TITLE
 JOURNAL
 BDBP/HMW Drosophila EST Project
 Unpublished (2001)
 Other_ESTs: Lp11963.3prime
 COMMENT
 Contact: Stapleton, M.

BDBP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
 hit genomic AE003593; arm:3L [20799555,21112669]
 estimated-cyto:78A2-78C7: 04/11/2001
 Plate: Lp.119 row: F column: 3
 High quality sequence stop: 558
 POLYA-No.

FEATURES

SOURCE

Location/Qualifiers
 1. 656
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /clone="Lp11963"
 /sex="male and female"
 /dev_stage="larvae-pupae"
 /lab_host="DHS-alpha"
 /clone_lib="Lp Drosophila melanogaster larval-early pupal
 potB7"
 /note="Organ: whole body; Vector: potB7; Site: 1: EcorI;
 Site: 2: XhoI; Sized fractionated cDNAs were directly
 ligated into potB7. Plasmid cDNA library."

ORIGIN

Query Match 12.5%; Score 645; DB 1; Length 656;
 Best Local Similarity 99.8%; Pred. No. 2.3e-177; Indels 1; Gaps 1;
 Matches 656; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 531 CTTATGATGAGGAGCCACAGAGTATGTTGCGCTATATATACCAACAGTTTGTGCA 590
 DB 1 CTTATGATGAGGAGCCACAGAGTATGTTGCGCTATATATACCAACAGTTTGTGCA 60
 OY 591 CTTCTCGAGATCCCAAGTTATGTGAGGAGCGGCGCATTCAAACGCTATGAAAAATC 650
 DB 61 CTTCTCGAGATCCCAAGTTATGTGAGGAGCGGCGCATTCAAACGCTATGAAAAATC 120
 OY 651 TACAGCATGTGAGGAGATCGATTGGCCCAAGCTCGGTGATGAGAGATGTTCTGCGC 710
 DB 121 TACAGCATGTGAGGAGATCGATTGGCCCAAGCTCGGTGATGAGAGATGTTCTGCGC 180
 OY 711 TCGAATTTGGCTATGTTGGAGCAAAAGTTTCAGCAAGTCAAAACAGAGGCTTACTGCTA 770
 DB 181 TCGAATTTGGCTATGTTGGAGCAAAAGTTTCAGCAAGTCAAAACAGAGGCTTACTGCTA 240
 OY 771 CTTGAGCCCTTAAAAACAAGATGGAATGAGTGGGCTTGGAGAGAGCCGACATATT 830
 DB 241 CTTGAGCCCTTAAAAACAAGATGGAATGAGTGGGCTTGGAGAGAGCCGACATATT 300

QY 831 GGGTTGAGGAGCGACCCAGGATGATTAGCGGCCACTACACTCGCCGTTTGTCTCA 890
 DB 301 GGGTTGAGGAGCGACCCAGGATGATTAGCGGCCACTACACTCGCCGTTTGTCTCA 360
 QY 891 TCACTGGGCCCCAAACCAATGTAACGATGACCCGGTATGCGCGCGCGGTACCATG 950
 DB 361 TCACCTGGGCCCCAAACCAATGTAACGATGACCCGGTATGCGCGCGCGGTACCATG 420
 QY 951 GAATCTTTGCAATCTAGCTTTGAGGTGGTCCCGCAATTTGAACATCTTCCAGCTAAGAC 1010
 DB 421 GAATC-TTCGATCTAGCTTTGAGGTGGTCCCGCAATTTGAACATCTTCCAGCTAAGAC 479
 QY 1011 ATGAGCATATCTCAAGCAAGTACTAGTATCATCAGTGATTAATAAAGCAGACTGGAG 1070
 DB 480 ATGAGCATATCTCAAGCAAGTACTAGTATCATCAGTGATTAATAAAGCAGACTGGAG 539
 QY 1071 AAACGTGTGATGCTCTCAAGAAATCAGAGGCTTGTCTCATCTTCCAGCTTACACTCAG 1130
 DB 540 AAACGTGTGATGCTCTCAAGAAATCAGAGGCTTGTCTCATCTTCCAGCTTACACTCAG 599
 QY 1131 CCGCAGTTTGTCTGCTGTAAGCTAAAGAAATTTGTCTTAAAGCTTCTGAGCATCTC 1187
 DB 600 CCGCAGTTTGTCTGCTGTAAGCTAAAGAAATTTGTCTTAAAGCTTCTGAGCATCTC 656

RESULT 10
 AG977859/c 642 bp DNA linear GSS 01-FEB-2005
 LOCUS Drosophila simulans DNA, clone: DSM1-021J17.F.fa, genomic survey
 DEFINITION
 ACCESSION AG977859
 VERSION AG977859.1 GI:58526752
 KEYWORDS GSS.

SOURCE Drosophila simulans
 ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1
 Hattori, M., Toyoda, A., Murakami, K., Kuroki, Y., Fujiyama, A.,
 Toshio, T. K. and Sakaki, Y.
 BAC end sequences of library DNBI
 Unpublished.
 2 (bases 1 to 642)
 Hattori, M.
 Direct Submission
 Submitted (19-JUN-2005) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
 (E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 Clones are derived from the BAC library DSM1
 For BAC library availability, please contact Masa-Toshi Yamamoto
 (yamamoto@kit.jp).
 Submitted (30-11-2004) by Masahira Hattori,
 RIKEN, Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: hattori@gsc.riken.jp, Tel: 81-45-503-9111,
 Fax: 81-45-503-9170)

COMMENT This work was done in collaboration with Yamamoto, M-T. Drosophila
 Genetic Resource Center
 Suga Ippongi-cho, Ukyo-ku, Kyoto 616-8354, Japan
 Tel: 81-75-873-2660 FAX: 81-75-861-0881
 PRIMERS
 Sequencing : F

LIBRARY : PKS150
 Vector : SacI
 R.Site 1 : SacI
 R.Site 2 : SacI
 Location/Qualifiers
 1. 642
 /organism="Drosophila simulans"
 /mol_type="genomic DNA"

ORIGIN
 /db_xref="taxon:7240"
 /clone="DSM1-021J17.F.fa"
 /clone_lib="DSM1 Drosophila BAC library"

Query Match 12.0%; Score 619.6; DB 10; Length 642;
 Best Local Similarity 97.8%; Pred. No. 6,7e-170;
 Matches 628; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 2040 CAGTCACAGCCAGGATGAGATTCACCTCGCCAGACAACTGAGGATCAGTACGGT 2099
 DB 642 CAGTCACAGCCAGGATGAGATTCACCTCGCCAGACAACTGAGGATCAGTACGGT 583
 QY 2100 GGTATTGTAATTAATCTATCTGATGAGGAGCTGAGGAGCCATACCCAAAGGCTCTCGGAAT 2159
 DB 582 GGTATTGTAATTAATCTATCTGATGAGGAGCTGAGGAGCCATACCCAAAGGCTCTCGGAAT 523
 QY 2160 CCCCAGACAGCCAGCTCTAGGGAAGAAGTCCAACTAGGTCAGTGTGGCTTGATG 2219
 DB 522 CCCCAGACAGCCAGCTCTAGGGAAGAAGTCCAACTAGGTCAGTGTGGCTTGATG 463
 QY 2220 AAACGAGATGATCTCTACAGGTGGGGGTCTGACGTACGCCCGAGAGAAACCA 2279
 DB 462 AAACGAGATGATCTCTACAGGTGGGGGTCTGACGTACGCCCGAGAGAAACCA 403
 QY 2280 GTAAACCATGCGCGCGGCGACGACTGCTGGCGCAATCCGTGAGACAGAACTACATTA 2339
 DB 402 GTAAACCATGCGCGCGGCGACGACTGCTGGCGCAATCCGTGAGACAGAACTACATTA 343
 QY 2340 GGGGTGGAGATGATGAGAACGACGATGTTTCCGGGGAATCACTAGGCGCGCGGA 2399
 DB 342 GGGGTGGAGATGATGAGAACGACGATGTTTCCGGGGAATCACTAGGCGCGCGGA 283
 QY 2400 ATGCGGATGAGTGAAGAGCTCATGGAAGTGTGATGATCATGACCTCCAGGCGC 2459
 DB 282 ATGCGGATGAGTGAAGAGAGCTCATGGAAGTGTGATGATCATGACCTCCAGGCGC 223
 QY 2460 AGTTCTGTGTGTTTCAAGACGATCTTTTGAATCCGAGTACATGAGGTAATGAAC 2519
 DB 222 AGTTCTGTGTGTTTCAAGACGATCTTTTGAATCCGAGTACATGAGGTAATGAAC 163
 QY 2520 TACTCATTAAGGGAGGACCAACCGCTTGACTGAGAGAGAGGCGGCATTGAC 2579
 DB 162 TACTCATTAAGGGAGGACCAACCGCTTGACTGAGAGAGAGGCGGCATTGAC 103
 QY 2580 GACATGAGACGATTAATTCATGTTCTGCGCATGACGACATTTGTTGTAAGAGATGGC 2639
 DB 102 GACATGAGACGATTAATTCATGTTCTGCGCATGACGACATTTGTTGTAAGAGATGGC 43
 QY 2640 CTGATCAGGCTTCAACACAGTATCTGGCCGATGGAAGAGAGCTC 2681
 DB 42 CTGATCAGGCTTCAACACAGTATCTGGCCGATGGAAGAGAGCTC 1

RESULT 11
 BG634723 606 bp mRNA EST 23-APR-2001
 LOCUS AT31034.5prime AT Drosophila melanogaster adult testes p07B7
 DEFINITION Drosophila melanogaster cDNA clone AT31034.5, mRNA sequence.
 ACCESSION BG634723
 VERSION BG634723.1 GI:13762260
 KEYWORDS EST.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 606)
 Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E.,
 Bernan, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,
 Farfan, D., Frisoe, E., George, R., Gonzalez, M., Garin, H., Harris, N.,
 Li, P., Liao, G., Miranda, A., Miera, S., Mungall, C. J., Nuno, J.,
 Pacle, J., Paragae, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,

TITLE Lewis, S.E., Celniker, S. and Rubin, G.M.
JOURNAL BDGP/HMI At Drosophila EST Project
COMMENT Unpublished (2000)
Contact: Stapleton, M.

BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AEO03593: arm:3L [20799555, 21112669]
estimated-cyto:78A2-78C7: 04/09/2001
Plate: AT.310 Row: C Column: 10
High quality sequence stop: 551.
Location/Qualifiers

FEATURES
source

1..606
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT31034"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_id="AT Drosophila melanogaster adult testes POTB7"
/note="Organ: ADULT testes; Vector: POTB7; Site_1: EcORI;
Site_2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into POTB7. Plasmid cDNA library."

ORIGIN

Query Match 11.7%; Score 604.4; DB 2; Length 606;
Best Local Similarity 99.8%; Pred. No. 1.9e-165;
Matches 605; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2887 AGATTGGAGACGCTACAGGTGTGCACGAATATTTCCACGACGTTGACCTTAAG 2946
1 AGATTGGAGACGCTACAGGTGTGCACGAATATTTCCACGACGTTGACCTTAAG 60
2947 AACTATTGAATCATATCGGATTTACTCAAAGCCCAACCAAGAGCGGCAATGCCA 3006
61 AACTATTGAATCATATCGGATTTACTCAAAGCCCAACCAAGAGCGGCAATGCCA 120
3007 TTCTGCGCTCTTACGATCTGCGCAATACATCTTAAGACGACCTTTCCACG 3066
121 TTCTGCGCTCTTACGATCTGCGCAATACATCTTAAGACGACCTTTCCACG 180
3067 ACCAGAGCAGGCTGCGAGCGGAGCGTCTTAAGCTGCGCCACGCTGCGGCGATCGA 3126
181 ACCAGAGCAGGCTGCGAGCGGAGCGTCTTAAGCTGCGCCACGCTGCGGCGATCGA 240
3127 AGTCAGAGAGTGGCTGCCAGGCGCAGAGTGGCTTATGAGCCCTGTATTAACCTGAATA 3186
241 AGTCAGAGAGTGGCTGCCAGGCGCAGAGTGGCTTATGAGCCCTGTATTAACCTGAATA 300
3187 CCGCGCAATGACCTTTTACCTGCGCGACCTGCGCAAAAGGTATTCAGAGCTTGCAGAT 3246
301 CCGCGCAATGACCTTTTACCTGCGCGACCTGCGCAAAAGGTATTCAGAGCTTGCAGAT 360
3247 CCGCATCATTCGACATGAGGCGGCAAAAGGTGCAATTCGGGTGCCAATTCGC 3306
361 CCGCATCATTCGACATGAGGCGGCAAAAGGTGCAATTCGGGTGCCAATTCGC 420
3307 CTAGTAGCTCTTCATTTAGAGCAGTAGCTCCCAAGCTTTTGCAGAACTCCCTGTGGGTC 3366
421 CTAGTAGCTCTTCATTTAGAGCAGTAGCTCCCAAGCTTTTGCAGAACTCCCTGTGGGTC 480
3367 CATTTGCTCTGCTTGAAGCCACACCACTTATGACATGACCTTACTAGTCCAGCT 3426
481 CATTTGCTCTGCTTGAAGCCACACCACTTATGACATGACCTTACTAGTCCAGCT 540
3427 CCGGAGAGCTTCCGAGAGAGAGAGCTCTTTTCTCGAGCTGACATTCAGACACA 3486

DB 541 CCGGAGAGCTTCCGAGAGAGAGAGCTCTTTTCTCGAGCTGACATTCAGACACA 600
QY 3487 ACATTC 3492
DB 601 ACATTC 606

RESULT 12
BF493211 626 bp mRNA linear EST 19-APR-2001
LOCUS AT01524.5prime AT Drosophila melanogaster adult testes POTB7
DEFINITION Drosophila melanogaster cDNA clone AT01524 5, mRNA sequence.
BP493211
ACCESSION BP493211.2 GI:13683797
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
1 (bases 1 to 626)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E.,
Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,
Farfan, D., Fries, B., George, R., Gonzalez, M., Guarin, H., Harris, N.,
Li, P., Liao, G., Miranda, A., Miera, S., Mungall, C. J., Nuno, J.,
Pacled, J., Pargatz, V., Park, S., Phuanavong, S., Wan, K., Yu, C.,
Lewis, S.E., Celniker, S. and Rubin, G.M.
BDGP/HMI At Drosophila EST Project
Unpublished (2000)
On Dec 6, 2000 this sequence version replaced gi:11576512.
Contact: Stapleton, M.

TITLE
JOURNAL
COMMENT
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AEO03593: arm:3L [20799555, 21112669]
estimated-cyto:78A2-78C7: 04/07/2001
Plate: AT.15 Row: B Column: 12
High quality sequence stop: 500.
Location/Qualifiers

FEATURES
source

1..626
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT01524"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_id="AT Drosophila melanogaster adult testes POTB7"
/note="Organ: ADULT testes; Vector: POTB7; Site_1: EcORI;
Site_2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into POTB7. Plasmid cDNA library."

ORIGIN

Query Match 11.6%; Score 597; DB 2; Length 626;
Best Local Similarity 100.0%; Pred. No. 3e-163;
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

14 GCCGCTCCGAATTTGTTGCGCAGAACGAAAGCAATTAACATTTGCAAGTGGCCTA 73
30 GCCGCTCCGAATTTGTTGCGCAGAACGAAAGCAATTAACATTTGCAAGTGGCCTA 89
74 TCGGAAGCCCGACGACCTGATGCTTATTCAGCAAAATGCGCAAGCGGACATGCTGT 133
90 TCGGAAGCCCGACGACCTGATGCTTATTCAGCAAAATGCGCAAGCGGACATGCTGT 149
134 GAAGTACAGCTCGCGAGAGATCTGTGACATTCCTTAGCAGACGACAACTCAATGT 193

```

Db      150 GAAGGATACAGCTGCGGAGATCTGTGACATTCCTTAGGAGACACAAACCAATTCCT 209
QY      194 CTCACGACGATGGATTCCTTATTTAGACGTTTATGATGACATGAGCGGACGACCTT 253
Db      210 CTCACGACGATGGATTCCTTATTTAGACGTTTATGATGACATGAGCGGACGACCTT 269
QY      254 TAAGATTGACAAAGTCCCTGGAGCGTTCTCGAGCTTAATAAGCGATTGGACGCA 313
Db      270 TAAGATTGACAAAGTCCCTGGAGCGTTCTCGAGCTTAATAAGCGATTGGACGCA 329
QY      314 TTTTATGATACACAGCTTCCCTTTCGACATGATGATCGGCTGGAGACGACG 373
Db      330 TTTTATGATACACAGCTTCCCTTTCGACATGATGATCGGCTGGAGACGACG 389
QY      374 GGACACAGTCCGAGAGGCGCAACTTCTGTCGCGGACCTCATGAGACAGATGCT 433
Db      390 GGACACAGTCCGAGAGGCGCAACTTCTGTCGCGGACCTCATGAGACAGATGCT 449
QY      434 TCCGCCCGAGCGCTGATGACAGCTGCGCACTAGCTGCTTCAGACAGACGCA 493
Db      450 TCCGCCCGAGCGCTGATGACAGCTGCGCACTAGCTGCTTCAGACAGACGCA 509
QY      494 GGTGCGCGAGAGATTCCTTCAAGCATTTGAAAGCTTTCATGATGACGACGCA 553
Db      510 GGTGCGCGAGAGATTCCTTCAAGCATTTGAAAGCTTTCATGATGACGACGCA 569
QY      554 GCTTAGTGTTCGCGCTATATACACAGTTTGTGACCTTCGAGATCCCACT 610
Db      570 GCTTAGTGTTCGCGCTATATACACAGTTTGTGACCTTCGAGATCCCACT 626

```

```

RESULT 13
LOCUS   BF491691               615 bp    mRNA    linear    EST 23-APR-2001
DEFINITION   AT28490.5 prine AT Drosophila melanogaster adult testes potB7
ACCESSION   BF491691
VERSION     BF491691.2 GI:13756851
KEYWORDS    EST.
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
REFERENCE   1 (bases 1 to 615)
AUTHORS    Stapleton, M., Brocks, P., Hong, L., Agbayani, A., Baxter, E.,
            Bernan, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,
            Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N.,
            Li, P., Liao, G., Miranda, A., Miera, S., Mungall, C. J., Nuno, J.,
            Pacleb, J., Paragay, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,
            Lewis, S. E., Ceinkner, S. and Rubin, G. M.
            BDGP/HMI AT Drosophila EST Project
            Unpublished (2000)
            On Dec 6, 2000 this sequence version replaced gi:11574992.
            Contact: Stapleton, M.
            BDGP
TITLE      JOURNAL
COMMENT    Lawrence Berkeley National Lab
            One Cyclotron Rd. Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
            hit genomic AEO03593: atm:3L [20799555,21112669]
            estimated-cyto:78A2-78C7: 04/09/2001
            Plate: AT.284 row: H column: 6
            High quality sequence scop: 596.
            Location/Qualifiers
            1..615
            /organism="Drosophila melanogaster"
            /mol_type="mRNA"
            /db_xref="taxon:7227"
            /clone="AT28490"
            /sex="male"
            /dev_stage="0-3 day old Ore-R males"

```

```

/lab host="Plates AT.10-AT.120: DHS-alpha. Plates
AT.121-AT.319: DHS-alpha Tona"
/clone lib="AT Drosophila melanogaster adult testes potB7"
/note="Organ: ADULT testes; Vector: the testis library was made
Site 2: Xhol; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into potB7. Plasmid cDNA library."

```

```

Query Match      11.6%; Score 596; DB 2; Length 615;
Best Local Similarity 100.0%; Pred. No. 5.8e-163;
Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      14 GCCGTCGGAATTGTTGTCGCGAGAGAGAGCAATAGCATTTGACGATGCGCTA 73
Db      20 GCCGTCGGAATTGTTGTCGCGAGAGAGAGCAATAGCATTTGACGATGCGCTA 79
QY      74 TCGAAGCCAGCGACCTGATGCTTTATCCAGCAATGCCCAAGCGGACATGCTGT 133
Db      80 TCGAAGCCAGCGACCTGATGCTTTATCCAGCAATGCCCAAGCGGACATGCTGT 139
QY      134 GAAGTACAGTCCGCGAGAGATCTGTGTACATTCCTTTAGCGACGACAACTCAATTGT 193
Db      140 GAAGTACAGTCCGCGAGAGATCTGTGTACATTCCTTTAGCGACGACAACTCAATTGT 199
QY      194 CTGACGACATGGGATTCCTTATTTGACGTTTATGATGCGATGCGTGAAGGACGACCTT 253
Db      200 CTGACGACATGGGATTCCTTATTTGACGTTTATGATGCGATGCGTGAAGGACGACCTT 259
QY      254 TAAGATTGACAAAGTCCCTGAGAGCGCTTCTGAGCTAATAAGCGATTGGAGCGCA 313
Db      260 TAAGATTGACAAAGTCCCTGAGAGCGCTTCTGAGCTAATAAGCGATTGGAGCGCA 319
QY      314 TTTTATGATACACAGCTTACCGTACCGTTCGACATGTGATGATGCGTGGAGACAGAG 373
Db      320 TTTTATGATACACAGCTTACCGTACCGTTCGACATGTGATGATGCGTGGAGACAGAG 379
QY      374 GGACACAGTCCGAGAGAGGCGCAACTTCTGTCGCGACCTCATGAGACAGAGTGTCT 433
Db      380 GGACACAGTCCGAGAGAGGCGCAACTTCTGTCGCGACCTCATGAGACAGAGTGTCT 439
QY      434 TCCGCCCGAGCGCTGATGACAGCTGCGCACTAGCTGCTTCAGACAGACGCA 493
Db      440 TCCGCCCGAGCGCTGATGACAGCTGCGCACTAGCTGCTTCAGACAGACGCA 499
QY      494 GGTGCGCGAGAGATTCCTTCAAGCATTTGAAAGCTTTCATGATGACGACGCA 553
Db      500 GGTGCGCGAGAGATTCCTTCAAGCATTTGAAAGCTTTCATGATGACGACGCA 559
QY      554 GCTTAGTGTTCGCGCTATATACACAGTTTGTGACCTTCGAGATCCCACT 609
Db      560 GCTTAGTGTTCGCGCTATATACACAGTTTGTGACCTTCGAGATCCCACT 615

```

```

RESULT 14
LOCUS   CK657543               640 bp    mRNA    linear    EST 30-JAN-2004
DEFINITION   LP23927.3 prine LP Drosophila melanogaster larval-early pupal pot2
ACCESSION   CK657543
VERSION     CK657543.1 GI:41401068
KEYWORDS    EST.
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
REFERENCE   1 (bases 1 to 640)
AUTHORS    Harvey, D., Brocks, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
            Lewis, S. and Rubin, G. M.
            BDGP/HMI Drosophila EST Project
TITLE      JOURNAL

```

JOURNAL
COMMENT

Unpublished (2001)
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or more T residues at the beginning of the sequence, this clone probably contains an inverted insert. The resulting Poly-T sequence has been removed.
Plate: LP.239 row: C column: 3
High quality sequence stop: 517.
Location/Qualifiers

FEATURES

1. 640
/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="LP23927"

/sex="male and female"

/dev_stage="larvae-pupae"

/lab_host="DHS-alpha"

/clone_lib="LP Drosophila melanogaster larval-early pupal"

/note="Organ: whole body; Vector: pOT2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pOT2. Plasmid cDNA library."

ORIGIN

Query Match 11.5%; Score 591; DB 7; Length 640;

Best Local Similarity 100.0%; Pred. No. 1.7e-161;

Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 14 GCCGCTCCGAATTGTTGCGCAGAGAGAGAGATGACATTGCAACGATGCGCTA 73
Db 50 GCCGCTCCGAATTGTTGCGCAGAGAGAGAGATGACATTGCAACGATGCGCTA 109
QY 74 TCGGAAGCCCGAGCGCTGGATGCTTTATCCAGAAATGCCAGAGCGGACATCGTGT 133
Db 110 TCGGAAGCCCGAGCGCTGGATGCTTTATCCAGAAATGCCAGAGCGGACATCGTGT 169
QY 134 GAAGGTACAGCTCGCGAGAGATCTGTGATCTCTTACGAGACACAACTCAATTGT 193
Db 170 GAAGGTACAGCTCGCGAGAGATCTGTGATCTCTTACGAGACACAACTCAATTGT 229
QY 194 CTGCAAGCAATGAGATTCCTTATTTGACCGGTTGATGCTGACGCGGAGCGACCT 253
Db 230 CTGCAAGCAATGAGATTCCTTATTTGACCGGTTGATGCTGACGCGGAGCGACCT 289
QY 254 TAAAGTTGACAAAAGTCCCTGAGGCGTTCTCGAGCTAATTAAGCAATTGGGACGCA 313
Db 290 TAAAGTTGACAAAAGTCCCTGAGGCGTTCTCGAGCTAATTAAGCAATTGGGACGCA 349
QY 314 TTTTATGATACACGCGTACCGTTCTGACACATGTGATCGCTGGAGAGACGAG 373
Db 350 TTTTATGATACACGCGTACCGTTCTGACACATGTGATCGCTGGAGAGACGAG 409
QY 374 GAGACAGTCCGCGAGAGAGCGCACTTGTCTGCGCGACCTCATGAGACAGAGTGT 433
Db 410 GAGACAGTCCGCGAGAGAGCGCACTTGTCTGCGCGACCTCATGAGACAGAGTGT 469
QY 434 TCGGCCCCGAGCGCTGATGACAAAGTGGCACTAGCTGTTCAAGACAAAGACCCAA 493
Db 470 TCGGCCCCGAGCGCTGATGACAAAGTGGCACTAGCTGTTCAAGACAAAGACCCAA 529
QY 494 GGGGCGGAGAGAGTCTCTTACAGAGATTGGAAGCGCTTTCATAGAGTACGACCGACA 553
Db 530 GGGGCGGAGAGAGTCTCTTACAGAGATTGGAAGCGCTTTCATAGAGTACGACCGACA 589
QY 554 GCTTAGTGTTCGGCTATATATACCAAGTTTGACCTTCTGGAGATCC 604
Db 590 GCTTAGTGTTCGGCTATATATACCAAGTTTGACCTTCTGGAGATCC 640

```

RESULT 15

BI171453

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE
JOURNAL
COMMENT

BDGP/HMT RE Drosophila EST Project
Contact: Stapleton, M.

BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AEO03593: arm:3L [20799555,21112669]
estimated-cyto.78A2-78C7: 04/12/2001
Plate: RE.132 row: C column: 10
High quality sequence stop: 526.
Location/Qualifiers

FEATURES

source

1. 631
/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="RE13234"

/sex="male and female"

/dev_stage="0-24 hours mixed stage embryonic"

/lab_host="DHS-alpha Tona"

/clone_lib="RE Drosophila melanogaster normalized Embryo pRc-1"

/note="Organ: embryo; Vector: pRc1; Site 1: XhoI; Site 2: BamHI; Library was kindly generated by Piero Carninci at the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."

ORIGIN

Query Match 11.4%; Score 587.2; DB 2; Length 631;

Best Local Similarity 97.6%; Pred. No. 2.3e-160;

Matches 615; Conservative 0; Mismatches 12; Indels 3; Gaps 2;

```

QY 2683 CTCAGACCACTTAATAGCGCTTCGACATGTTCCGAAAGATGTTCAATGATACACA 2742
Db 5 CTCAGACCACTTAATAGCGCTTCGACATGTTCCGAAAGATGTTCAATGATACACA 64
QY 2743 CCAAGGTACTCACTGTTTGTGACACAGGTACCGAACTGATTTGTTCAATGCGAAGC 2802
Db 65 CCAAGGTACTCACTGTTTGTGACACAGGTACCGAACTGATTTGTTCAATGCGAAGC 124
QY 2803 AAACCTTCAAGAAAGTCTTTCATCTTTGTTGACCGGTTGTTCAAAATTTGGGACCG 2862
Db 125 AA--CTTCAAGAAAGTCTTTCATCTTTGTTGACCGGTTGTTCAAAATTTGGGACCG 181
QY 2863 ACTTACTGAATTCGATGACAGAGATTGGAAGCCTACAGGTGATCGAGATATT 2922
Db 182 ACTTACTGAATTCGATGACAGAGATTGGAAGCCTACAGGTGATCGAGATATT 241
QY 2923 TCCCAAGCAGTTGAGCTTAAGAACTATTGAATCATATCGAATTCTACAAAGC 2982

```

```
Db      242 TCCCCAGCGAGTTGACGCTTAAGAACTATTAGATCATATCGGATTCTACTCAAAAGC 301
QY      2983 CAACTACCAAGAGCGCGCATTCGCGCTTCCTTAGGATCTGGCCAATACATACT 3042
Db      302 CAACTACCAAGAGCGCGCATTCGCGCTTCCTTAGGATCTGGCCAATACATACT 361
QY      3043 GTAAAAGCAGCGACTTTCCAGCGACCAAGCGCGCTGCGAGCGGACGGTCTTAAAGC 3102
Db      362 GCAAAAAGCAGCGACTTTCCAGCGACCAAGCGCGCTGCGAGCGGACGGTCTTAAAGC 421
QY      3103 TGGCCAGCTGGGGGGGATCAGAACTGATGAGCTGGCGCTCCAGGCGCAGAGCTGCC 3162
Db      422 TGGCCAGCTGGGGGGGATCAGAACTGATGAGCTGGCGCTCCAGGCGCAGAGCTGCC 481
QY      3163 TAGTGCCCTGTATAACCTGAATACCCCGCAATGACCCCTTTACTGGCCGACCTGCCAA 3222
Db      482 TAGTGCCCTGTATAACCTGAATACCCCGCAATGACCCCTTTACTGNCGACCTGCCAA 541
QY      3223 AGGTATATCAGACTCTGCCGATCCTGCATCCATTGCAATGAGGCGGCAAAAGCCAAA 3282
Db      542 AGGTATATCAGACTCTGCCGATCCTGCATCCATTGCAATGAGGCGGCAAAAGCCAAA 601
QY      3283 GTTGCAATTGGGGTGGCAATTGGCCTTAGTA 3312
Db      602 GTTGCAATTGGGGTGGCAATTGGCCTTAGTA 631
```

Search completed: November 30, 2005, 00:40:44
Job time : 12430 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 00:05:42 ; Search time 1882 Seconds
(without alignments)
18219.896 Million cell updates/sec

Title: US-10-030-850-1

Perfect score: 5145
Sequence: 1 gaattcgagccgagccgctc.....ggattcaaatcaataaaaa 5145

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_21:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*
- 14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5145	100.0	5145	4 AAF24991	Aa124991 Nucleotid
2	149.8	2.9	3153	10 ADB62173	Adb62173 Human CDN
3	149.8	2.9	7888	13 ADB83507	Adb83507 Human mul
4	143.4	2.8	6245	3 AAc64201	Aac64201 Human pol
5	141.8	2.8	2524	10 ADB62363	Adb62363 Human CDN
6	116.6	2.3	1908	10 ADB61959	Adb61959 Human CDN
7	116.6	2.3	5614	10 ADB79874	Adb79874 Rat putat
8	116.6	2.3	5614	14 ADX07257	Adx07257 Cyc1in-de
9	116.6	2.3	6445	4 Aa158443	Aa158443 Human pol
10	116.6	2.3	6445	5 AAd98654	Aad98654 DNA encod
11	116.6	2.3	6445	9 ADB48414	Adb48414 Novel hum
12	116.6	2.3	6487	5 Aa158442	Aa158442 Human pol
13	116.6	2.3	6487	5 ADQ98653	Adq98653 DNA encod
14	116.6	2.3	6487	9 ADB48413	Adb48413 Novel hum
15	116.6	2.3	6491	4 Aa160228	Aa160228 Human pol
16	116.6	2.3	6491	4 Aa160229	Aa160229 Human pol
17	115	2.2	1467	6 Aa160229	Aa160229 Human pol
18	115	2.2	6491	4 Aa160229	Aa160229 Human pol
19	98.4	1.9	6951	10 ADG32852	Adg32852 Human DNA

C	20	92	1.8	608	5 ABV59210	Abv59210 Human pro
C	21	80.2	1.6	695	10 ADF53926	Adf53926 Human con
C	22	52.8	1.0	615	5 ABV59086	Abv59086 Human pro
C	23	52.6	1.0	2000	8 ADA71938	Ada71938 Rice gene
C	24	50.6	1.0	486	5 AAs70766	Aas70766 DNA encod
C	25	49.8	1.0	371	4 AAK63080	Aak63080 Human imm
C	26	49.4	1.0	2000	11 AC137108	Ac137108 Rice stre
C	27	48.4	0.9	2000	11 AC135887	Ac135887 Rice stre
C	28	46	0.9	548	13 AD078622	Ad078622 Novel can
C	29	45.4	0.9	2000	8 ADA71938	Ada71938 Rice gene
C	30	45.2	0.9	561	5 AAs70767	Aas70767 DNA encod
C	31	44.2	0.9	1187	6 ABQ41869	Abq41869 Oligonuc1
C	32	44.2	0.9	1187	6 ABQ41868	Abq41868 Oligonuc1
C	33	42.8	0.8	2645	2 AAV62158	Aav62158 HSV-2 str
C	34	42.8	0.8	2841	2 AAV62146	Aav62146 HSV-2 str
C	35	42.8	0.8	3471	2 AAT93650	Aat93650 Herpes ai
C	36	42.8	0.8	3472	2 AAQ48496	Aaq48496 Glycoprot
C	37	42.8	0.8	117213	2 AAV62176	Aav62176 HSV-2 str
C	38	42	0.8	5047	11 ADN95651	Adn95651 Human BEC
C	39	42	0.8	8210	6 ABK1380	Abk1380 Signal tr
C	40	42	0.8	8210	6 ABL70331	Ab170331 Chemicali
C	41	42	0.8	8210	6 AAs61282	Aas61282 Human gen
C	42	41.8	0.8	528	6 ABQ29879	Abq29879 Oligonuc1
C	43	41.8	0.8	528	6 ABQ29878	Abq29878 Oligonuc1
C	44	41.8	0.8	1566	13 ADX09689	Adx09689 Plant ful
C	45	41.6	0.8	1863	13 ADX59482	Adx59482 Plant ful

ALIGNMENTS

RESULT 1	AAf24991	standard; cDNA; 5145 BP.
ID	AAf24991	standard; cDNA; 5145 BP.
AC	AAf24991;	
XX	30-APR-2001	(first entry)
DT	Nucleotide sequence of a Drosophila polypeptide designated orbit.	
DE	Orbit; chromosome segregation; mitotic spindle; microtubule; mitosis;	
XX	polyploid cell; microtubule binding; nucleation; motor protein; cancer;	
KW	leukemia cell; solid tumour; ss.	
KM	Drosophila sp.	
XX	OS	
XX	Key	Location/Qualifiers
FT	CDS	66..4544
FT		/*tag= a
FT		/product= "Orbit"
XX	MO200104295-A1.	
XX	18-JAN-2001.	
PD	11-JUL-2000; 2000WO-GB002662.	
XX	13-JUL-1999; 99GB-00016402.	
PR	24-DEC-1999; 99GB-00030707.	
XX	(UYDU-) UNIV DUNDEE.	
PA	Avides MDC, Deak P, Glover DM;	
XX	WPI; 2001-138345/14.	
XX	P-PSDB; AAB31597.	
DR	Novel Drosophila orbit protein and its human homolog, useful for treating	
XX	tumors and in assays for identifying substances capable of inhibiting	
PT	mitosis.	
PT	Claim 3; Page 84-85; 92pp; English.	
PS		

XX The present sequence encodes a Drosophila polypeptide designated orbit.
 CC The orbit protein is required for chromosome segregation. Orbit protein
 CC localises to mitotic spindles and binds microtubules. Orbit mutant
 CC phenotypes result in embryos exhibiting abnormal mitosis and polyploid
 CC cells. Orbit polypeptides are useful for inhibiting mitosis, and for
 CC identifying substances which affect orbit functions such as microtubule
 CC binding, microtubule organizing centre, nucleation activity and
 CC interactions with microtubule motor proteins. Tests for detecting or
 CC sequencing orbit, or its homologue, in a biological sample may be used to
 CC determine orbit sequences within cells in individuals who have or are
 CC suspected to have, an altered orbit gene sequence, for example within
 CC cancer cells including Leukaemia cells and solid tumours such as breast,
 CC ovary, lung, colon etc

Sequence 5145 BP; 1388 A; 1331 C; 1295 G; 1131 T; 0 U; 0 Other;

Query Match 100.0%; Score 5145; DB 4; Length 5145;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 5145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GAATTCGACGACGCGCTCCGAATTTGTTGCGGAGAACGAGGAGCATGACATTG 60
DB 1 GAATTCGACGACGCGCTCCGAATTTGTTGCGGAGAACGAGGAGCATGACATTG 60
QY 61 CAACGATGCTTATCGAAGCCGACCTGGATGGCTTTATCCAGCAATGCCCAAG 120
DB 61 CAACGATGCTTATCGAAGCCGACCTGGATGGCTTTATCCAGCAATGCCCAAG 120
QY 121 CGGACATGCTGTGAAGGTACAGCTGCGGAGATTTGGTACATTCTTACGACGACA 180
DB 121 CGGACATGCTGTGAAGGTACAGCTGCGGAGATTTGGTACATTCTTACGACGACA 180
QY 181 CAACTCAATTCTGCGACGACATGGGATTCCTTATGACGGTTTGAATGCCATGGCTGA 240
DB 181 CAACTCAATTCTGCGACGACATGGGATTCCTTATGACGGTTTGAATGCCATGGCTGA 240
QY 241 CGGAGACGACCTTTAAGTTGACAAAGTCCCTGGAGCGCTTTCGAGACTAATAAGC 300
DB 241 CGGAGACGACCTTTAAGTTGACAAAGTCCCTGGAGCGCTTTCGAGACTAATAAGC 300
QY 301 GATTGGGCGCATTTTAATGATACACGCTTACCGCTTTCGACATGATGATCGGC 360
DB 301 GATTGGGCGCATTTTAATGATACACGCTTACCGCTTTCGACATGATGATCGGC 360
QY 361 TGGGAGACGACGAGGACACAGTCCGAGAGAGCGCACTTCTGCGCGGACCTCATGG 420
DB 361 TGGGAGACGACGAGGACACAGTCCGAGAGAGCGCACTTCTGCGCGGACCTCATGG 420
QY 421 AGCAGAGAGTCTTCGCGCCGACGCGCTGATGACAGCTGCGCACTAGCTCTTCAAGC 480
DB 421 AGCAGAGAGTCTTCGCGCCGACGCGCTGATGACAGCTGCGCACTAGCTCTTCAAGC 480
QY 481 ACAAGAGCGCCAAAGTGCAGAGAGTTCCTTCAAGCATTTGTGAACGCTCTTCAAGT 540
DB 481 ACAAGAGCGCCAAAGTGCAGAGAGTTCCTTCAAGCATTTGTGAACGCTCTTCAAGT 540
QY 541 ACGGACCCGACGAGCTAGTGTTCGCGCTATATACACCAAGTTGTGCACTTCCGAG 600
DB 541 ACGGACCCGACGAGCTAGTGTTCGCGCTATATACACCAAGTTGTGCACTTCCGAG 600
QY 601 ATCCCAAGTTAATGTGAGGAGCGGCCATCAACGCTAGTGAATCTCAAGCATG 660
DB 601 ATCCCAAGTTAATGTGAGGAGCGGCCATCAACGCTAGTGAATCTCAAGCATG 660
QY 661 TAGGGGATGATTTGGGCCGACCTTCGCTGATGAGCATTTTCTGCTGGAATTGG 720
DB 661 TAGGGGATGATTTGGGCCGACCTTCGCTGATGAGCATTTTCTGCTGGAATTGG 720
QY 721 CTATGTTGAGCAAAAGTTGACACAGGTCAACAGAGGGGTCTTACTGCTTCAAGCC 780
DB 721 CTATGTTGAGCAAAAGTTGACACAGGTCAACAGAGGGGTCTTACTGCTTCAAGCC 780

```

```

QY 781 TTAACACAGATGCAATGAGTGGGCTTGGACGAGGCCGACATATGGGTTAGGG 840
DB 781 TTAACACAGATGCAATGAGTGGGCTTGGACGAGGCCGACATATGGGTTAGGG 840
QY 841 AGCGACCCACAGATGATTTAAGGCGCCACTACACTGGCCGTTTCCGATCAGTCC 900
DB 841 AGCGACCCACAGATGATTTAAGGCGCCACTACACTGGCCGTTTCCGATCAGTCC 900
QY 901 CAAACCCATGTGAACGATGACCGGTGATGCGGCGCGTAAACATGGAATCTTTCG 960
DB 901 CAAACCCATGTGAACGATGACCGGTGATGCGGCGCGTAAACATGGAATCTTTCG 960
QY 961 AATCTAGCTTTGAGGTGCTCCGCAATTGAACATCTTTCACGCTTAAGACATGACGATA 1020
DB 961 AATCTAGCTTTGAGGTGCTCCGCAATTGAACATCTTTCACGCTTAAGACATGACGATA 1020
QY 1021 TCTACACAGATGCTATGATGATCAATGATGATTAACACGACGACGAGAAACGTTGG 1080
DB 1021 TCTACACAGATGCTATGATGATCAATGATGATTAACACGACGACGAGAAACGTTGG 1080
QY 1081 ATGCTCTCAAGAAATCAGGCGATTGCTCATTTCTCAGCTATCAGCTCAGCGCAGTTTG 1140
DB 1081 ATGCTCTCAAGAAATCAGGCGATTGCTCATTTCTCAGCTATCAGCTCAGCGCAGTTTG 1140
QY 1141 TCGCTGTACAGCTTAAGGAATTTGCTTAAAGCTTCGTGACATCTTCAAGAGAACTAC 1200
DB 1141 TCGCTGTACAGCTTAAGGAATTTGCTTAAAGCTTCGTGACATCTTCAAGAGAACTAC 1200
QY 1201 GATCAGAGGTATCCGCGAGCGGCGATCAACATCGCTTACATGCTTAAGACGCTGAGAA 1260
DB 1201 GATCAGAGGTATCCGCGAGCGGCGATCAACATCGCTTACATGCTTAAGACGCTGAGAA 1260
QY 1261 ATAACTAGATGCTTCTGCTGAGACATTTTGAAGACCTGATTAATTAATACGAAACA 1320
DB 1261 ATAACTAGATGCTTCTGCTGAGACATTTTGAAGACCTGATTAATTAATACGAAACA 1320
QY 1321 GCGCGAAGTCAATGATCCGCTTCAACATAGCTCTGAAGTATATCTTAAGTATAC 1380
DB 1321 GCGCGAAGTCAATGATCCGCTTCAACATAGCTCTGAAGTATATCTTAAGTATAC 1380
QY 1381 ATGACCAAGATCTCTCAAGTATCAACAGACCTCTGAATCAATCAAAAGTCAAGGACA 1440
DB 1381 ATGACCAAGATCTCTCAAGTATCAACAGACCTCTGAATCAATCAAAAGTCAAGGACA 1440
QY 1441 TAAGTCCACACTGTGTGAGCTGATGTGTCTCTTTCGAGAGTGGCAGAGGCGT 1500
DB 1441 TAAGTCCACACTGTGTGAGCTGATGTGTCTCTTTCGAGAGTGGCAGAGGCGT 1500
QY 1501 TGAAGGAATGCAACGCTACTAAGGACACCTTAAACAAATCCATTGGGATGACACT 1560
DB 1501 TGAAGGAATGCAACGCTACTAAGGACACCTTAAACAAATCCATTGGGATGACACT 1560
QY 1561 GCGATGACGCGCGCATTTCCAGATACGCTATTTGGGCTTTCAGGCGTCACTTCCAGAGC 1620
DB 1561 GCGATGACGCGCGCATTTCCAGATACGCTATTTGGGCTTTCAGGCGTCACTTCCAGAGC 1620
QY 1621 TGGCGGATCAAAATATATGAAACATTAGACATAGCTGCGCAGCGGCAATTGAAAGGAAAC 1680
DB 1621 TGGCGGATCAAAATATATGAAACATTAGACATAGCTGCGCAGCGGCAATTGAAAGGAAAC 1680
QY 1681 GAGAGGGCGCGGAGAGAGGAACTGTACTGGGACTGGGACTGACCTGAAACGAGAC 1740
DB 1681 GAGAGGGCGCGGAGAGAGGAACTGTACTGGGACTGGGACTGACCTGAAACGAGAC 1740
QY 1741 GCACTGTATCCCGATTGGCCGACCACTTGAACCTTCCAAACCCCAAGCTTATGATGA 1800
DB 1741 GCACTGTATCCCGATTGGCCGACCACTTGAACCTTCCAAACCCCAAGCTTATGATGA 1800
QY 1801 GATCCATTTCAGCGGTGAGACATCGCGCTGCTCAACGAGCCAAAGTTAAGAGCGCAATATA 1860
DB 1801 GATCCATTTCAGCGGTGAGACATCGCGCTGCTCAACGAGCCAAAGTTAAGAGCGCAATATA 1860
QY 1861 CACTATATTCAGGCAAAAGAACCTTTAGAGACTAATTAATTCACACAGGATCATGATGA 1920

```


Db 1861 CACTATTATCCAGGAAAGAAACCTTTAGACCTAATAATTCCAACGAGATGATGA 1920
 Qy 1921 CAGGACGACGACCGAGTGGATCACTGCCAGACCTCGCCTGAATTCCTAATAGCGGTGCA 1980
 Db 1921 CAGGACGACGACCGAGTGGATCACTGCCAGACCTCGCCTGAATTCCTAATAGCGGTGCA 1980
 Qy 1981 CACGAGTACAAACGCGGGATCGTTTACACACGCGCCCGGGGACGGGACGAGATATCTC 2040
 Db 1981 CACGAGTACAAACGCGGGATCGTTTACACACGCGCCCGGGGACGGGACGAGATATCTC 2040
 Qy 2041 AGTCACAGCCAGATCGAGATCCACTCGCCAGACAAATAATGAGGATCATGTCGGTG 2100
 Db 2041 AGTCACAGCCAGATCGAGATCCACTCGCCAGACAAATAATGAGGATCATGTCGGTG 2100
 Qy 2101 GTATTGGTAACTTACTATCGTGGGGCCACTGGCGCCATACCAAAAAGGCTCCGGAATAC 2160
 Db 2101 GTATTGGTAACTTACTATCGTGGGGCCACTGGCGCCATACCAAAAAGGCTCCGGAATAC 2160
 Qy 2161 CCCGAGACACAGCCAGCTCTAGGGAAACGAGTCCAATAGTCAAGTGGTGGCTTTGATGA 2220
 Db 2161 CCCGAGACACAGCCAGCTCTAGGGAAACGAGTCCAATAGTCAAGTGGTGGCTTTGATGA 2220
 Qy 2221 AACGAGTATGTACTCTTACAGGTGGGGGTCTTGACGTACGCCGAGAGAAACACCCAG 2280
 Db 2221 AACGAGTATGTACTCTTACAGGTGGGGGTCTTGACGTACGCCGAGAGAAACACCCAG 2280
 Qy 2281 TAAAGCAATCGCGCGCGGACAGACTGTCGGGCAATCCGCTGAGAGAGAACTACATTAG 2340
 Db 2281 TAAAGCAATCGCGCGCGGACAGACTGTCGGGCAATCCGCTGAGAGAGAACTACATTAG 2340
 Qy 2341 GCGTCGAGATGATGAGACAGCCAGACTATGTTCCGGGGACTYACATGCGACGCGCGAA 2400
 Db 2341 GCGTCGAGATGATGAGACAGCCAGACTATGTTCCGGGGACTYACATGCGACGCGCGAA 2400
 Qy 2401 TCGCGATGGGATGAGAACTCATGAGACGTGATGAGTCTGATGACATCCACTCCGAGGCA 2460
 Db 2401 TCGCGATGGGATGAGAACTCATGAGACGTGATGAGTCTGATGACATCCACTCCGAGGCA 2460
 Qy 2461 GTTCGTGTGTTCAGAAAGATCTTTCAGTCCAGCTACACTAGAGGTAAATAATGAACT 2520
 Db 2461 GTTCGTGTGTTCAGAAAGATCTTTCAGTCCAGCTACACTAGAGGTAAATAATGAACT 2520
 Qy 2521 ACTCATTTAGCGGAGCCACACCCGCTTGAATGAGCAACGACGCGGCGCAATTGAGC 2580
 Db 2521 ACTCATTTAGCGGAGCCACACCCGCTTGAATGAGCAACGACGCGGCGCAATTGAGC 2580
 Qy 2581 ACATCGAGACGATTAATTCAGTTCTGCGCATCGACGATGGTCTGAAAGAAAGATGGCC 2640
 Db 2581 ACATCGAGACGATTAATTCAGTTCTGCGCATCGACGATGGTCTGAAAGAAAGATGGCC 2640
 Qy 2641 TCAATCAGCTTCAACAGTATCTGCGGATGAGAAAGAGCTCATCGACGAACTTAAT 2700
 Db 2641 TCAATCAGCTTCAACAGTATCTGCGGATGAGAAAGAGCTCATCGACGAACTTAAT 2700
 Qy 2701 GCGTTCTGACATGTTCCGCAAGATGTTATGATGATACACACACGAGGTACTCATCTGT 2760
 Db 2701 GCGTTCTGACATGTTCCGCAAGATGTTATGATGATACACACACGAGGTACTCATCTGT 2760
 Qy 2761 TTTTGGACACGGTCAACGAACTGATTCGTGGTCAATGAGAACTTCAAGAAATGGCT 2820
 Db 2761 TTTTGGACACGGTCAACGAACTGATTCGTGGTCAATGAGAACTTCAAGAAATGGCT 2820
 Qy 2821 CTTCATCTTGTGAGCGGCTTGTCAACAAATTGGGACCGATCTACCAATTCGATGC 2880
 Db 2821 CTTCATCTTGTGAGCGGCTTGTCAACAAATTGGGACCGATCTACCAATTCGATGC 2880
 Qy 2881 ACAGCAAGATTGGAAGACGCTACAGTGGTCCACGAAATATTTCCCAACGACAGTTGCAAC 2940
 Db 2881 ACAGCAAGATTGGAAGACGCTACAGTGGTCCACGAAATATTTCCCAACGACAGTTGCAAC 2940
 Qy 2941 TTAAGAACTATTAGATCAATCGGATTTCTAATAACGCACTACCAAGACCGCA 3000
 Db 2941 TTAAGAACTATTAGATCAATCGGATTTCTAATAACGCACTACCAAGACCGCA 3000

Db 2941 TTAAGAACTATTAGATCAATCGGATTTCTAATAACGCACTACCAAGACCGCA 3000
 Qy 3001 TCGCCATTCTGCGCTTCTTACGAGATCTGGCCAAATACATCTGTAATAAGACGCACTTTC 3060
 Db 3001 TCGCCATTCTGCGCTTCTTACGAGATCTGGCCAAATACATCTGTAATAAGACGCACTTTC 3060
 Qy 3061 CCAGCGACACAGGACGAGCGCGGAGCGGACCGGCTTAAGCTGAGCCAGCTGGCGGCGG 3120
 Db 3061 CCAGCGACACAGGACGAGCGCGGAGCGGACCGGCTTAAGCTGAGCCAGCTGGCGGCGG 3120
 Qy 3121 ATCAGAGTCGATGAGCTGCGCTCCAGGACAGAGCTGCTAGTGGCCCTGTATTAAC 3180
 Db 3121 ATCAGAGTCGATGAGCTGCGCTCCAGGACAGAGCTGCTAGTGGCCCTGTATTAAC 3180
 Qy 3181 TGAATACCCCGCAATGACCTTTTATCGGCCACCTGCGCAAAAGTATATAGAACTCTG 3240
 Db 3181 TGAATACCCCGCAATGACCTTTTATCGGCCACCTGCGCAAAAGTATATAGAACTCTG 3240
 Qy 3241 CCCGATCCTGATCATTTCCGACATGAGGCGGCAAGCAAGTTGCATTCGGGTGCA 3300
 Db 3241 CCCGATCCTGATCATTTCCGACATGAGGCGGCAAGCAAGTTGCATTCGGGTGCA 3300
 Qy 3301 ATTGCGCTAGTGTCTTCATTTAGAGAGTACAGTCCCAAGCTTTGCAAAATCCCTCTG 3360
 Db 3301 ATTGCGCTAGTGTCTTCATTTAGAGAGTACAGTCCCAAGCTTTGCAAAATCCCTCTG 3360
 Qy 3361 TGGGTCAATTTGCTCGCTTCAGAGCCACACACCAACTTAAGATCAGCTCTACTATGC 3420
 Db 3361 TGGGTCAATTTGCTCGCTTCAGAGCCACACACCAACTTAAGATCAGCTCTACTATGC 3420
 Qy 3421 CAGCCTCCCGGACCTTCCGTCGAGAGAGAGGCTCTTTTCCGAGGCTGACATTC 3480
 Db 3421 CAGCCTCCCGGACCTTCCGTCGAGAGAGAGGCTCTTTTCCGAGGCTGACATTC 3480
 Qy 3481 AGCAACACATTCAGAAAGCTCGAGAGAGATCCGCACTGCTTGGCGGTGATACAGA 3540
 Db 3481 AGCAACACATTCAGAAAGCTCGAGAGAGATCCGCACTGCTTGGCGGTGATACAGA 3540
 Qy 3541 CGGCGCTGGCGCCCAATGCTTCAATGAGCACTTGCAATTCAGATTCAGAGGCCAAACAG 3600
 Db 3541 CGGCGCTGGCGCCCAATGCTTCAATGAGCACTTGCAATTCAGATTCAGAGGCCAAACAG 3600
 Qy 3601 ATTGCTGCAATCCCTGCTTTCACATCTCAAGAGCAATCGTCGAGCACTACCCAGT 3660
 Db 3601 ATTGCTGCAATCCCTGCTTTCACATCTCAAGAGCAATCGTCGAGCACTACCCAGT 3660
 Qy 3661 CAATATACCTGAGTCAGCACAATGAGCTGATTAATCTGAGCGGGAAAGACCACTC 3720
 Db 3661 CAATATACCTGAGTCAGCACAATGAGCTGATTAATCTGAGCGGGAAAGACCACTC 3720
 Qy 3721 AGAAGCCCAAGTCAACCACTGACATGCAAGTGCATGATCAAGTCTGATTAATAGCTG 3780
 Db 3721 AGAAGCCCAAGTCAACCACTGACATGCAAGTGCATGATCAAGTCTGATTAATAGCTG 3780
 Qy 3781 AAAATGAGACTGTATATGCGCCAGCAACCTGATGGAAGGGAAGTGGTGGCT 3840
 Db 3781 AAAATGAGACTGTATATGCGCCAGCAACCTGATGGAAGGGAAGTGGTGGCT 3840
 Qy 3841 TGAAGCTTAAAGGATCAGCCCGTCAAGTGTCTTCAAGCTCATTAACCTGGGA 3900
 Db 3841 TGAAGCTTAAAGGATCAGCCCGTCAAGTGTCTTCAAGCTCATTAACCTGGGA 3900
 Qy 3901 TTTGATCAAGGCGGAAACTGTGAGCTGCCCAATAGCACTTAAATGATGATGCGAA 3960
 Db 3901 TTTGATCAAGGCGGAAACTGTGAGCTGCCCAATAGCACTTAAATGATGATGCGAA 3960
 Qy 3961 TGCCTCTTAACTTCTGAGAGCGGAGCATAGGAGTGTGATGCTGCGCTGCAAGTGC 4020
 Db 3961 TGCCTCTTAACTTCTGAGAGCGGAGCATAGGAGTGTGATGCTGCGCTGCAAGTGC 4020
 Qy 4021 TCAATAGATTATGAGAGCAACAAATGCTGCAACATGAGTGCATTTCTAGAGCTGA 4080
 Db 4021 TCAATAGATTATGAGAGCAACAAATGCTGCAACATGAGTGCATTTCTAGAGCTGA 4080

QY 4081 TTTTGTGTAAGATCATCCAGTGTCTATCAACACAGCAGAGAGCGTTTGGGGATATCGACT 4140
 DB 4081 TTTTGTGTAAGATCATCCAGTGTCTATCAACACAGCAGAGAGCGTTTGGGGATATCGACT 4140
 QY 4141 CGATGATATCAAGAGATACACACATCTCTGGGATCTGTCCATCAACATTTGCAATC 4200
 DB 4141 CGATGATATCAAGAGATACACACATCTCTGGGATCTGTCCATCAACATTTGCAATC 4200
 QY 4201 CGGTATATGCAACGGGTGAGTTTCCACAAAATTTGTGGCCATCAAGATCTGTGAGAG 4260
 DB 4201 CGGTATATGCAACGGGTGAGTTTCCACAAAATTTGTGGCCATCAAGATCTGTGAGAG 4260
 QY 4261 TGACCGAACACCATGTTTGGAGATCACAGAGCGCCATTGGAATTTGTTCCCAATC 4320
 DB 4261 TGACCGAACACCATGTTTGGAGATCACAGAGCGCCATTGGAATTTGTTCCCAATC 4320
 QY 4321 TGGGGCCATCGGGCGAGACACGCAATGATGTTTGGAAAGCTTCTTTCGATCG 4380
 DB 4321 TGGGGCCATCGGGCGAGACACGCAATGATGTTTGGAAAGCTTCTTTCGATCG 4380
 QY 4381 TCAAGCTGTACTTTGTTTGGGGCGAGAGAGTCAAGCCGAAGCTGTCACTGCTAAATC 4440
 DB 4381 TCAAGCTGTACTTTGTTTGGGGCGAGAGAGTCAAGCCGAAGCTGTCACTGCTAAATC 4440
 QY 4441 CCAGCAAGTTAGGCTCTCTCAACGTGTATCGAGAGACGCGAACTGCATCAATGCGG 4500
 DB 4441 CCAGCAAGTTAGGCTCTCTCAACGTGTATCGAGAGACGCGAACTGCATCAATGCGG 4500
 QY 4501 GAGGAAGCTTACAAAGAACTCTCCGGGATGTCGTCTGATGAGTGGGAGGCTTTAA 4560
 DB 4501 GAGGAAGCTTACAAAGAACTCTCCGGGATGTCGTCTGATGAGTGGGAGGCTTTAA 4560
 QY 4561 TAGGATTCCTGCTGCGCCACCAACAAGACAGAGCGGTTGCTTCCCTCGGCTG 4620
 DB 4561 TAGGATTCCTGCTGCGCCACCAACAAGACAGAGCGGTTGCTTCCCTCGGCTG 4620
 QY 4621 AGAAGAAAGTGAAGAGCGCGGACATTAATATATATATATTAATTAATTAATTA 4680
 DB 4621 AGAAGAAAGTGAAGAGCGCGGACATTAATATATATATATTAATTAATTAATTA 4680
 QY 4681 TAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4740
 DB 4681 TAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4740
 QY 4741 ATGCAAAATCGGCGCGAAGAAAGTCAAGGCGCATCGCCATCTAATGTGAGCAAAA 4800
 DB 4741 ATGCAAAATCGGCGCGAAGAAAGTCAAGGCGCATCGCCATCTAATGTGAGCAAAA 4800
 QY 4801 CTATTTATACATAACGGGAATAAGCGAGTAATCCGAAAGTGTACTTAAATTTGAG 4860
 DB 4801 CTATTTATACATAACGGGAATAAGCGAGTAATCCGAAAGTGTACTTAAATTTGAG 4860
 QY 4861 CCAACTCGGCAATCTCTCTCAACATCAAGTGTGACCTCTTTTCAACAGCTCTAATA 4920
 DB 4861 CCAACTCGGCAATCTCTCTCAACATCAAGTGTGACCTCTTTTCAACAGCTCTAATA 4920
 QY 4921 TTAAGCGGATTTGTTTGAATTTGTAAGCTATATACCGCTCTTATGAAAGTCAAGC 4980
 DB 4921 TTAAGCGGATTTGTTTGAATTTGTAAGCTATATACCGCTCTTATGAAAGTCAAGC 4980
 QY 4981 TGTGCTTTTCTACCTTTGTTTGAAGCGAATTTGTAATTTGTTTGGCTCAACTA 5040
 DB 4981 TGTGCTTTTCTACCTTTGTTTGAAGCGAATTTGTAATTTGTTTGGCTCAACTA 5040
 QY 5041 AATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5100
 DB 5041 AATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5100
 QY 5101 AATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5145
 DB 5101 AATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5145

RESULT 2
 ADB62173
 ID ADB62173 standard; cDNA; 3153 BP.
 XX
 AC ADB62173;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Human cDNA encoding clone CTONG20032930.
 KW Human; ss. gene; pharmaceutical; diagnostic; gene therapy;
 KW tissue regeneration; cell regeneration; membrane protein;
 KW signal transduction-related protein; transcription-related protein;
 KW osteoporosis; neurological disease; cancer; tumour.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 352..3102
 FT /*tag= a
 FT /product= "Clone CTONG20032930 protein"
 XX
 PN EPI308459-A2.
 XX
 PD 07-MAY-2003.
 XX
 PF 28-MAR-2002; 2002EP-00007401.
 XX
 PR 05-NOV-2001; 2001JP-00379298.
 PR 25-JAN-2002; 2002US-00350978.
 XX
 PA (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Iisogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamehika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Nasuno Y;
 XX
 DR WPI; 2003-450961/43.
 DR P-PSDB; ADB64143.
 XX
 PT New polynucleotides and polypeptides, useful for developing a diagnostic
 PT marker or medicines for regulation of their expression and activity, or
 PT as targets of gene therapy.
 XX
 PS Claim 1; Page; 222pp; English.
 XX
 CC The invention discloses a polynucleotide comprising a sequence selected
 CC from 1970 fully defined nucleotide sequences which encode novel
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
 CC or its partial peptide, an antibody binding to the polypeptide or peptide
 CC of the polynucleotide, immunologically assaying the polypeptide or
 CC peptide of the polynucleotide by contacting the polypeptide or peptide
 CC with the antibody of the encoded protein, and observing the binding
 CC between the two, a transformant carrying the polynucleotide in an
 CC expressible manner and an antisense polynucleotide. The oligonucleotide
 CC is useful as a primer for synthesising the polynucleotide, or as a probe
 CC for detecting the polynucleotide. The polynucleotides and encoded
 CC proteins are useful as pharmaceutical agents and many disease-related
 CC genes may be included in them, for developing a diagnostic marker or
 CC medicines for regulation of their expression and activity, or as targets
 CC of gene therapy. The genes are involved in tissue and/or cell
 CC regeneration. Membrane proteins, signal transduction-related proteins,
 CC transcription-related proteins, disease-related proteins and genes
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
 CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
 CC the activity or expression of the encoded protein to treat diseases. The
 CC sequence presented is a cDNA of the invention. Note: Some of the sequence
 CC data for this patent is not represented in the printed specification, but
 CC is based on sequence information supplied by the European Patent Office.
 XX
 SQ Sequence 3153 BP; 883 A; 701 C; 767 G; 802 T; 0 U; 0 Other;

Query Match 2.9%; Score 149.8; DB 10; Length 3153;
 Best Local Similarity 47.1%; Pred. No. 8.3e-32;
 Matches 685; Conservative 0; Mismatches 732; Indels 36; Gaps 6;

QY 232 CATGCTGACGGGACCACTTTAAGATTGACAAAAGTCCCTGGAGGCGTTCTGGAGC 291
 DB 521 CTGGGATGAATCTGACCAATTAACAAGGTGTTCTGCTGGGACATGCACTGTCGGCC 580
 QY 232 TAAATAGGATTTGGGACCGATTTTAATGACATACCGGCTACCGTTCCGACATGGA 351
 DB 581 TGGTAACTCCGCTGACAGATCGGTTCAAGCGAGATCGGACAGTGGCTGCGCAAGCTTA 640
 QY 352 TCGATCGCTGGGAGACACAGGAGACACAGTCCGCGAGAGGCGCAACTTCTGCTCGCG 411
 DB 641 TAGACAGATAGAGATGCTTAAGACTCTGTGAGGAGACAGGACCAAACTCTGCTGCTAA 700
 QY 412 ACCTGATGAGACACAGATGCTTCCGCGCCAGCGCTGATGCAACAGCTGGCCACTAGCT 471
 DB 701 AGATCATGATCAAGCTGCTAAATCCCACTAGCTATGAGACAGAAATGCT-----TGGAG 754
 QY 472 GCTTCAAGACAAAGACCGCAAGGTCCGAGAGAGTTCCTCAGAGATTTGAAAGCTC 531
 DB 755 GCTTCAAAACAAGAAATTTCCGTACTCGAAGGACATCTGCTCTGCTTATAGCAAC 814
 QY 532 TTCTGATGATGAGGACCCAGCACTTATGTTGCGCTCTATATACCAACAGTTTGTGAC 591
 DB 815 TCAATGCTCTGAGACACAGACTTTTAACACTAAGCAAGATTTGTCACATATATGCACT 874
 QY 592 TTCTGAGAGATCCACAGTTAAATGTAGAGGAGCGGCGCATCCAAACGTTAGTGAATCT 651
 DB 875 TACTTGGAGATCAAAACAGCGGTTCGAGATGACAGCAATAAACAGCTTAGTGAATTT 934
 QY 652 ACAAGCATTTAGGGATTCATGCGCCAGACCTCCGTGCAATGACAGATTTCCGCT 711
 DB 935 ACGACATGTAGAGAACGTGTAGGAGGAGATCTCGTAA--AAAGATTCCACAGCT 991
 QY 712 CGAAATTTGGCTATGTTGAGACAAAAGTTGACACAGGCTCAACAGAGGCTTACTGCTAC 771
 DB 992 CCCGTTGATGTAATTTTTCACAAATTTGATGAAATCCAGAAATTTGAAACATGATAC 1051
 QY 772 CTTCAAGCCCTTAAAAACAGAAATGGCAATGAGTGGCTTGAACGAGGCC-GACAAATAT 830
 DB 1052 AATTCGAATATATAAAATTTGACAGATGAAGATCTGTGATGTTAACAACCTTCCCT 1111
 QY 831 GGGTTAGAGGAGCGACCCACAGAGATTTAAGCGGCCACTACCTCGGCCCTTTCGTCA 890
 DB 1112 CTGCTAATTTCAATATCCAAAGGCTCCACCAAGTTCTCGAGAAACGTTGGAATGGAA 1171
 QY 891 -----TCACTGCGCCAAAACCAATGTGAACGATGTGACCGGTGATG 933
 DB 1172 CAACCCGCGGCTGTTGATTCACCTTGATTCAGTCCAAAGCTTCAGTGCAGAAAGAGAG 1231
 QY 934 CCGGCGCCGTAACCATGAAATCTTTGAAATCTAGCTTTGAGGTGTCGCGCAATTGACA 993
 DB 1232 CTGGTCTGTGATGAAGAGATTTTATTAAGCAATTTGATGATGATCTGATGATCAGA 1291
 QY 994 TCTTCAAGCTTAAGACATGAGATATCTACAAGCAAGTACTAGTATCATAGAGAA 1053
 DB 1292 TTTATTCAGCCGACCTTGAAGAAATCAATAAAACAAATTTAGGAAATTTATCTGATG 1351
 QY 1054 AAAAGCAGACTGGGAGAAACGTGTGATGCTCTCAAGAAAGATCAGGCAATTGCTCATTC 1113
 DB 1352 ACAAGCATGTTGGAGCAGAGATTAATGCTTAAAGAAATTAATCTTTACTTT-- 1409
 QY 1114 TCAGTATACACTCAGCCGAGTTTGTGCTGTACAGCTAAAGAAATGTGCTTAAGCT 1173
 DB 1410 ---GGCTGTGCTGTGATGATGATTAATCTTTCAACATTTGCTCTTTTGGATGGAG 1465
 QY 1174 TGTGACATCTCTCAAGAGGAACTACGATCAAGGTGATCCCGGAGGCGTGTGATCAACA 1233
 DB 1466 CTTTAAACTCTGTGTAAGGACCTCGGTCTCAGGATGATGCGGAGGCTTGTATACGCT 1525

QY 1234 TCGCTACATGCTTAAGACGCTGAGAAATTAACATGATGCTTCTGCTGAGCAATTTGG 1293
 DB 1526 TGGGCACTGCTCATCAAGTTCGGGGAATTAAGTTGACCATGAGCTGAGCCATTAATGC 1585
 QY 1294 AGCACTGATTAATTTAATACAGAACGCGGAAGTCAATTGCAATCCGCTTCCACAAATAG 1353
 DB 1586 CAACATCTTTAATTTAATTTTCAAAACAGTCCCAAAATTTATGCCACATCTGCTGTAG 1645
 QY 1354 CTCTGAAGTATATGATTAATGATTAACATGACCAACCAAGCTCTCAAGATCTTACACAGCA 1413
 DB 1646 CTGTTAGTTAATTAATTCGACACACATCCCTAGTTAATCTGTCTATTAACAAAGCA 1705
 QY 1414 CTCTGATCAATCAAAAGTTCGAAGACATTAAGTCCACTGTGAGAGTGTAGTGGCTGC 1473
 DB 1706 --ACTGTACCTCTTAAGCTGTGCAAGTAAAGCGCTGTTTGAATTTTGAATTTGC 1762
 QY 1474 TCTTCAAGAGTGGGACAGAAAGGCTTGAAGAAATGACCAAGTACTTAAGGACACCT 1533
 DB 1763 TTTTACAGAAATGACAGACATTTCACTAGAACACATATCAATTAAGCTGAAACAA 1822
 QY 1534 TAAATAATTCATTTGGCGATGACAGATGACAGCGCGCATTCAGATACGCTTAT 1593
 DB 1823 TAAAGAGGAATACATGATGCTGATTCGAAAGCAAGATAGAAAGCCGAAATGTTACT 1882
 QY 1594 GGGCTTCAAGCGCTCACTTTCCAGAGCTGGGGGATCAATAATATGAAACATTAACATAG 1653
 DB 1883 GGGGTTTCCACAGTCACTTCCAGAGAGAGAGACACTTGTACCAACCTTGGAGTCT 1942
 QY 1654 CTGCGCCAGCGCGC 1666
 DB 1943 CTTACAGAAAGC 1955

RESULT 3
 ADR83507
 ID ADR83507 standard; DNA; 7888 BP.
 XX
 AC ADR83507;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Human multiple asters 1 DNA, target gene of miRNA.
 XX
 DE human; ds; miRNA; microRNA; ontogenesis; cell therapy; cancer;
 KW immune disease; nerve disorder; amyotrophic lateral sclerosis;
 KW Parkinson's disease; Alzheimer's disease; inflammatory disease;
 KW siRNA silencing precursor; cytosolic; immunosuppressive; nootropic;
 KW neuroprotective; antiinflammatory; immunotherapy; multiple asters 1.
 XX
 OS Homo sapiens.
 XX
 PN WO2004076622-A2.
 XX
 PD 10-SEP-2004.
 XX
 PF 10-FEB-2004; 2004WO-JP001433.
 XX
 PR 10-FEB-2003; 2003US-0445829P.
 XX
 PA (NABD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
 XX
 PI Taira K, Kawasaki H;
 XX
 DR MPI; 2004-653393/63.
 XX
 PT Modulating expression of a target gene in a cell, for treating cancer, an
 PT immune disease, or a nerve disorder, comprises introducing into the cell
 PT a polynucleotide that forms a duplex region with an mRNA transcribed from
 PT the target gene.
 XX
 PS Claim 9; SEQ ID NO 409; 865bp; English.
 XX
 CC This invention relates to a novel method for modulating the expression of

CC a target gene in a cell. Specifically, it refers to the introduction into
 CC a cell of a polynucleotide that forms a duplex region with an mRNA
 CC transcribed from the target gene, where the duplex region comprises a
 CC mammalian miRNA target region i.e. a non-coding microRNA (miRNA) that
 CC regulates mRNA at a post-transcriptional level. The present invention
 CC describes a method for controlling the onset of a mammal, function of a
 CC mammalian cell, differentiation of a mammalian cell or viability of a
 CC mammalian cell in the post-transcriptional phase, which comprises
 CC introducing a plasmid vector comprising a promoter and nucleic acid
 CC molecule expressing an miRNA or siRNA silencing precursor to the mRNA.
 CC Accordingly, it provides a cell therapy method for treating cancer,
 CC immune disease, nerve disorder (e.g. amyotrophic lateral sclerosis,
 CC Parkinson's disease, or Alzheimer's disease) or an inflammatory disease
 CC by introducing into the cell the miRNA, siRNA silencing precursor to the
 CC miRNA or the plasmid vector. As such, they can be developed into
 CC pharmaceutical compositions that exhibit cytostatic, immunosuppressive,
 CC neurotropic, neuroprotective and anti-inflammatory activities and hence can
 CC be used for immunotherapy. This polynucleotide sequence is a human target
 CC gene whose expression is modulated by miRNAs of the invention.

XX Sequence 7888 BP; 2129 A; 1876 C; 1858 G; 2025 T; 0 U; 0 Other;

Query Match 2.9%; Score 149.8; DB 13; Length 7888;
 Best Local Similarity 47.1%; Pred. No. 1.5e-31;
 Matches 685; Conservative 0; Mismatches 732; Indels 36; Gaps 6;

QY 232 CATGCTGACGGGACGACCTTTAAGATTGCAAAAATCCCTGGAGCGTTCTCGAGC 291
 DB 335 CTGGGTGTAAGTCTGATGCAATTAAGAGTGTCTGCTGGGATGACATCTGTCGCC 394
 QY 292 TAAATAAAGGATGGGACGCAATTTAATGCAATACCGGCTACCGTTGCAATGGA 351
 DB 395 TGGTGAACCCGCTGAGGATCGGTTCAGAGCGCAAGTGGCAAGTCTGCAAGTCAA 454
 QY 352 TCGATCGCTGGGAGACAGGAGCAAGTCCGAGAGAGCGCAATTTGCTGGCGCG 411
 DB 455 TAGACAGCTGAGAGATCTTAAAGCTGTGAGGAGACAGACCAATCTGCTGCTAA 514
 QY 412 ACCATGAGAGACAGAGTGTCTCCGCCAGCGCTGATGCAAGCTGGCCACTACT 471
 DB 515 AGATCATGATGATCAAGCTCTAATCCCAAGTATGAGACAGATGT-----TGGAG 568
 QY 472 GCTTCAAGCAAGAACCCAGAGTGGCGGAGAGACTTCTTACAGCATTTGTAACGTC 531
 DB 569 GCTTCAAGCAAGAAATTTCCGTACTCGAAGAGGATCTGCTCTTATACCAAC 628
 QY 532 TTCAATGATAGGAGCCAGAGCTTAGTGTGGGTATATACCAATTTGTGAC 591
 DB 629 TCAATGCTTGGAGACAGACTTTAACAATAAGAGATTGTGCAATATATGCACT 688
 QY 592 TTCTGAGATCCACAGTATATGAGGAGGCGGCATCAACCGTAGTGAATCT 651
 DB 689 TACTTGGAGATCCAAACAGCGATGTCAGATGCAATTAACGCTTAGTGAATTT 748
 QY 652 ACAAGCATGAGGGGATGATGGCCCAAGCTTCGTCGATGAGCATGTTCTGCT 711
 DB 749 ACAAGCATGAGGAGAGCGTGTGAGGCGAGATCTCAGTAA---AAGATTGCGACAGT 805
 QY 712 CGAATTTGGCTATGTTGAGCAAAAGTTCGACCGAGTCAACAGAGGCTTACTGCTAC 771
 DB 806 CCGGTTGAATGTAATTTTTCACAAATTTGATGAAGTCAAAATCTGAAACATGATAC 865
 QY 772 CTTCAGCCCTTAAACACAGATGAGATGAGTGGCTTGGAGAGGCC-GACAAATTT 830
 DB 866 AATCTGCAAAATGATTAATTTTTCAGCATGAAGATTTGTGATGTAACAGACTTCT 925
 QY 831 GGGTTGAGGAGGACCAACAGATGATTAAGCGGCACATCACTGGCCGTTTCTGCA 890
 DB 926 CTGCTAGTTTACATCATCAAGGCTCCACCAAGTTCTCGAGAAACCTTGAATGGAA 985
 QY 891 -----TCACTGGCGCCAAACCAATGTAAGCATGATGACCGGTATG 933
 DB 986 CCACCCCGCGGCTTGTTCATCACTTGGATTCAGTCTTTCAGCTGCAAAAGAGAG 1045

QY 934 CCGGCGCGGTACCACTGAAATCTTTCGAAATCTGAGTGTGTCGCCAATTTGACA 993
 DB 1046 CTGGGTGTTGATGAAGAGATTTTATTAACATTTGATGATGTACCTGTATGACGA 1105
 QY 994 TCTTCCACGTTAAGCATGAGATGATATCTAACAGCAAGTACATGATCATCGATA 1053
 DB 1106 TTTATTTCCACCGAGACCTTGAAGAAATCTCAAAACAAATTTAGGAAATTTATCTATG 1165
 QY 1054 AAAACGAGACTGGGAGAAAGTGTGATGCTCTCAAGAAATCAGGGCATTTGCTCATTC 1113
 DB 1166 ACAAGCATGATTTGGAGACAGAGTAAATCTCTAATAAAGATTAGTCTTTACTTT-- 1223
 QY 1114 TCAAGTATCACTCAAGCCGCACTTGTGCTGTGACGTTAAAGAAATTTGCTTAAGCT 1173
 DB 1224 ----GGCTGTGCTGTGATGATGATTAATCTTTTCAACATTTTGCGCTTTTGGATGAG 1279
 QY 1174 TCGTGACATCTTCAGAGAGAACTACATCAAGGATCCGAGAGCGGTGATCAACA 1233
 DB 1280 CTTTAAACTCTCTGCTTAAGGACCTGCGGTCTAGAGTATGCGGAGCTTGTATACGT 1339
 QY 1234 TCGCTACATGCTTAAAGACGCTGAGAAATTAATAGATGCTTCTGCTGAGCATTTTGG 1293
 DB 1340 TGGGGCATCTGTATCATGATCTTGGGAAATAGTTGACATGAGCTGAAGCATTTATGC 1399
 QY 1294 AGCACTGATTAATTTAATACAGAACGCGGAAAGTCAATGCAATCCGTTTCAATAG 1353
 DB 1400 CAATATCTTTAATTTAATTTCAAAACAGTCCAAATTTATGGCCACATCTGTGTGTAG 1459
 QY 1354 CTGTAAGTATATCATTAAGTATACATGACCAAAAGCTGCTCAAGATCTACACAGCA 1413
 DB 1460 CTGTAGGTTAATTTATTTGCGACACACATCCCTAGTTAATCTGTATACAAAGCA 1519
 QY 1414 CTGTAATCAATCAAAATCGAAGACATTAAGTTCACATGTGTGAGCTGATGTCTGC 1473
 DB 1520 ---ACTGTACTCTTAAGTCTGTGCAATTTGAAAGGCGCTTTTGAATTTTATGATTC 1576
 QY 1474 TCTTCAGAGATGCGACAGAGCGTTGAAAGAAATGCCACCGTACTTAAGGACACT 1533
 DB 1577 TTTTACAAAGATGCGACACATTTCACTAGAACAGACATATGATTTAGCTGAAACAA 1636
 QY 1534 TAAATAAATCCATTTGGATGAGACATGCGATGACGCGCCATTCAGATACGCTATT 1593
 DB 1637 TAAAGAAAGGAAATCATGATGTGATTTCCAGACAAAGATAGAACCAAAATGTTACT 1656
 QY 1594 GGGCTTTCAGAGCTCACTTTCCAGAGCTGCGGATCAATATATATGAAACATTAACATAG 1653
 DB 1697 GGGTTTCCACAGTCACTTTGAGCAGAGAAACAGAGCACTTTACACACCTTGAATCT 1756
 QY 1654 CTGCCAGCGCGC 1666
 DB 1757 CTTACCGAAGAGC 1769

RESULT 4
 AAC64201
 ID AAC64201 standard; cDNA; 6245 BP.
 XX
 AC AAC64201;
 XX
 DT 21-FEB-2001 (first entry)
 XX
 DE Human polliosis-associated gene 373 cDNA, SEQ ID NO:1.
 XX
 KW Human; polliosis-associated gene 373; IgE; immunoglobulin E;
 KW cedar pollen allergy; T-cell; reduced expression; detection; diagnosis;
 KW drug screening; allergic disease; ss.
 XX
 OS Homo sapiens.
 XX
 MO200065046-A1.
 XX
 PD 02-NOV-2000.

XX 26-APR-2000; 2000WO-JP002730.
 PF 27-APR-1999; 994P-00120489.
 XX (GENO-) GENOX RES INC.
 XX Nagasu T, Sugita Y, Kashiwabara T, Oshida T, Obayashi M, Gunji S;
 PI Obayashi I, Imai Y, Yoshida N, Ogawa K, Matsui K;
 XX MPI: 2000-687339/67.
 DR P-PSDB; AAB29634.
 PT Polinosis-associated gene 373 undergoing significantly low expression in
 PT subjects with high cedar pollen-specific immunoglobulin-E levels. useful
 PT in diagnosis of allergic diseases and screening drug candidates.
 PS Claim 2; Page 41-57; 80pp; Japanese.

CC The invention relates to the human polinosis-associated gene 373 which
 CC exhibits significantly reduced expression in the T-cells of individuals
 CC with high cedar pollen-specific IgE (immunoglobulin E) levels. The gene
 CC was isolated from T-cells from individuals allergic to cedar pollen using
 CC the differential display method. The invention also relates also relates
 CC to the protein encoded by polinosis-associated gene 373; expression constructs and
 CC host cells comprising polinosis-associated gene 373 nucleic acids;
 CC polinosis-associated gene 373 primers and probes; antibodies against the
 CC protein encoded by the gene; methods of detection of polinosis-
 CC associated gene 373 nucleic acids; and a method of diagnosis of allergic
 CC diseases via the detection of polinosis-associated gene 373 nucleic
 CC acids. The invention additionally encompasses methods of screening drug
 CC candidates for the treatment of allergic disease by measuring the
 CC expression of polinosis-associated gene 373 in pollen antigen-stimulated
 CC T-cells in the presence of a test compound relative to a control.
 CC Polinosis-associated gene 373 is useful in the diagnosis of allergic
 CC diseases and in the screening of drug candidates for the treatment of
 CC such diseases. The present sequence represents human polinosis-
 CC associated gene 373 cDNA

XX Sequence 6245 BP; 1861 A; 1237 C; 1345 G; 1801 T; 0 U; 1 Other;

Query Match 2.8%; Score 143.4; DB 3; Length 6245;
 Best Local Similarity 46.2%; Pred. No. 9.7e-30;
 Matches 769; Conservative 0; Mismatches 851; Indels 45; Gaps 7;

QY 77 GAAGCCCAAGCACTGAGTGGCTTTATCGACAAGTCCCAAGGCGGAGCATGCTGTGA 136
 DB 1 GGAGCCCTCCAGATGAGTACTTCTGCGCCAGGTGACAGAAAGAGCTCGGCGCGC 60
 QY 137 GGTACAGCTCGCCGAGATCTGGTGAATTCCTTACGCG--ACGACACAACTCAATTGT 193
 DB 61 GCTGCAAGTCTGCGCAGAGCTCTCTGCTACCTTGGCGCCCGGCGCATCTCGGACCT 120
 QY 194 CTGCAAGCAATGAGATCTCT-----TATTGACGCTTGATGCCATGCTGACGGG 244
 DB 121 GGAGGAGACCTGGCGCCCTGAGCAAGACAGTGCAGCGCTCACCGCGCTGGGTGCTTC 180
 QY 245 CAGCCACTTAAATGACAAAGTCCCTGAGGGGCTCTCGGAGCTAATAAAGGATT 304
 DB 181 GAGCAACTACCGGGTATCAATTAATGAGATGAAATTTTAAGTCTTGTGTGACAGATT 240
 QY 305 GGGCAGCAGATTTAATGATACACAGGCTACCTTCTGCCACATGATGATGATGCGCTGG 364
 DB 241 ATCAACAGCTTTAATTCCTATGTAGCAATGTTATTTAGTATTTAATGACAGATGG 300
 QY 365 AGACAGCAGGAGACAGTCCGAGAGGCGCAATTTCTGCTGGCGGACCTCATGAGCA 424
 DB 301 AGATGCCAAGACAGAGTTCAGATGAAGCTCAGACTCTGATATTGAAGTTAATGATGA 360
 QY 425 CAGAGTCTTCCCGCCAGCGCTGATCGACAGAGCTGGCCACTAGGCTGCTTCAAGCA 484
 DB 361 AGTAG-----CACCACTATATGACATTGGGAGCAGTTGGCTTGTGTTTAAACACA 414

QY 485 GAAGCCCAAGTGGCGGAGAGTTCCTTACAGCATTTGTAAGCGCTTCAATGATACGG 544
 DB 415 GAATTTTGCATCGAAGAGCGTGTCTGTCTTATTTGAACCTTAAACATTTTGG 474
 QY 545 CACCAGCAGCTTAAATGTCGCTTATATACCAACAGTTTGTGACCTTCTCGAGATCC 604
 DB 475 GCTTCAGCACAATGATCAGACCAATTAATGATTCATTTGTTATTCCTGTTTGGAGACTC 534
 QY 605 CACAGTTAATGTGAGGAGGCGGCCATCAACCGTATGGAATCTACAGCATGTAG 664
 DB 535 CAACAGTCAAGTGAAGATGCTGCAATTTGGCTATATGAGATTATTAACATGTGG 594
 QY 665 GATGATTTGGGCCAGACCTCCGTCGATGAGAGATGCTCTGCGC-----TCGAATTTGG 720
 DB 595 AGAAAAGTGAAGTGAATCTTTATTAAGAGAGAAATTCCTCGTCAATTAAGAAATGAT 654
 QY 721 CTATGTTGAGCAAAAGTTGACCAAGTCAACAGAGAGGCTTACTGCTACCTTGACGCC 780
 DB 655 ATTTCGCAAAATTTGATGATGAGTGAAGTTCAGGCGGTATGATTTGAGTGTGCAAGA 714
 QY 781 TTAACAACAGATGAGCAATGAGTGAAGTGGCTTGACAGAGCGGACAAATTTGGCTTAGGG 840
 DB 715 TAAAGCTTGATGATGAAGATCAATGATGATGAATAGGCGCATCATGATGATGATGAT 774
 QY 841 AGCAGCCACAGATGATTTAAGCGGCACTACCTGCGCGCTTGC----- 887
 DB 775 CTTCAAGGTTCTTGACCTTAAACATTCGGAATCTTCCAACTGCAAGTCAAGAGACCTGG 834
 QY 888 -TCATCACTGCGCCCAAAACCAATGTGAGATGACCGGTGATGCGCGCGCTTAC 946
 DB 835 TTCAAGCAGTGGCCCTTAAGTGTGAGAGTCTTCTTAAGAGAGAGGTGCTGAGAGCTTCA 894
 QY 947 CATGGAATCTTTGCAATCTTGAATCTTGAATGAGTGTGCGCAATTTAAATCTTCAAGCTTA 1006
 DB 895 TGAAGATGATTTTAAAGCTTTTACAGATGCTCTTCTTATTCGATTTATTTCTAGTCG 954
 QY 1007 GCACTGAGCATATCTACAGCAAGTACTGTATCATCATGATTAATAACGACAGCTG 1066
 DB 955 AGAACTCGAAGAAATTAATTAATTAATGAAGAAATTTGTCAGATGATTAACATGACTG 1014
 QY 1067 GGAGAAAGCTGTGATGCTCTCAAGAGATGACGAGCATTTGCTCATTTCTAGCTATCACAC 1126
 DB 1015 GATATGAGTGTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
 QY 1127 TCAAGCGAGTGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1186
 DB 1069 TGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1128
 QY 1187 CAAGAGGAACTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1246
 DB 1129 AGCTAAGATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1188
 QY 1247 TAAAGCGTGAAGAAATTAATAGATGCTTCTGCTGAGACATTTTGAACCTGATTTAA 1306
 DB 1189 AAGCTTTTGGAAACAAAGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1248
 QY 1249 TCTGCTCCCAATAGTGAAGAAAGTATGAGCAACTTCTGATGATGATGATGATGATGATGAT 1308
 DB 1307 TTTAATACAGAAACAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1366
 QY 1367 CATTAAGTATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1426
 DB 1309 CATTGGGATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1365
 QY 1427 AAAGTGAAGACATTAAGTTCACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1486
 DB 1366 AAATTAAGTTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1425
 QY 1487 GCAAGCAGAGGCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1546
 DB 1426 GCAAGCTATTTATTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1485
 QY 1547 TGGCATGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1606

DB 1486 TCATATATCTGACGCTGAGGCGGAGTGAGCGCAAGAAAGACATCATGGCTTTAGAA 1545
 QY 1607 TCATCTTCAGAGCTGGCGGATCAATATATAGAACATTAGCATAGGCGCCAGCGGC 1666
 DB 1546 CCATCTTCTGTGTGAAGCTGAAACATTATATATTCCTTGAGCCATCTTATCAGAAAG 1605
 QY 1667 ATTAGAAAGGAACAGAGGCGGCGGAGAGAGAGAACTGGTAC 1711
 DB 1606 TCTTCAAACTTACTTAAAGAGTTCTGGCAGGTGTACATCTCTTCC 1650

RESULT 5

ADB62363
 ID ADB62363 standard; cDNA; 2524 BP.

XX ADB62363;

XX 04-DEC-2003 (first entry)

XX Human cDNA encoding clone FCBBF30105860.

XX Human; ss: gene; pharmaceutical; diagnostic; gene therapy;
 tissue regeneration; cell regeneration; membrane protein;
 signal transduction-related protein; transcription-related protein;
 osteoporosis; neurological disease; cancer; tumour.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 355..2524
 FT /tag= a
 FT /partial
 FT /product= "Clone FCBBF30105860 protein"

XX EPI308459-A2.

XX 07-MAY-2003.

XX 28-MAR-2002; 2002EP-00007401.

XX 05-NOV-2001; 2001JP-00379298.
 PR 25-JAN-2002; 2002US-00350978.

XX (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isegai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI: 2003-450961/43.
 DR P-PSDB; ADB64333.

XX New polynucleotides and polypeptides, useful for developing a diagnostic
 marker or medicines for regulation of their expression and activity, or
 as targets of gene therapy.

XX Claim 1; Page: 222pp; English.

XX The invention discloses a polynucleotide comprising a sequence selected
 CC from 1970 fully defined nucleotide sequences which encode novel
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
 CC or its partial peptide, an antibody binding to the polypeptide or peptide
 CC of the polynucleotide, immunologically assaying the polypeptide or
 CC peptide of the polynucleotide by contacting the polypeptide or peptide
 CC with the antibody of the encoded protein, and observing the binding
 CC between the two, a transformant carrying the polynucleotide in an
 CC expressible manner and an antisense polynucleotide. The oligonucleotide
 CC is useful as a primer for synthesizing the polynucleotide, or as a probe
 CC for detecting the polynucleotide. The polynucleotides and encoded
 CC proteins are useful as pharmaceutical agents and many disease-related
 CC genes may be included in them, for developing a diagnostic marker or

CC medicines for regulation of their expression and activity, or as targets
 CC of gene therapy. The genes are involved in tissue and/or cell
 CC regeneration. Membrane proteins, signal transduction-related proteins,
 CC transcription-related proteins, disease-related proteins and genes
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
 CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
 CC the activity or expression of the encoded protein to treat diseases. The
 CC sequence presented is a cDNA of the invention. Note: Some of the sequence
 CC data for this patent is not represented in the printed specification, but
 CC is based on sequence information supplied by the European Patent Office.
 XX

SO Sequence 2524 BP; 650 A; 590 C; 671 G; 613 T; 0 U; 0 Other;

Query Match 2.8%; Score 141.8; DB 10; Length 2524;
 Best Local Similarity 46.1%; Pred. No. 1.6e-29;
 Matches 768; Conservative 0; Mismatches 852; Indels 45; Gaps 7;

QY 77 GAAGCCAGAGCAGCTTGATGCTTTATCCAGCAAAATCCCAAGCGGACATGCTGTGAA 136
 DB 357 GGAACCCCGGAGCATGAGTGAATCTTGGCCGACAGTGCAGAGAGACGTGGCGCG 416
 QY 137 GGTACAGCTGCGCGAGGATCTGGTGAACATTCCTTAGCG---ACGACAAACCTCAATTGT 193
 DB 417 GCTGCAAGTGGCCAGAGCTCCTGCTTACCTTTGGCGCCCGCCCATCTGCACT 476
 QY 194 CTGACCGACATGAGATTCTCT-----TATTGACGTTTGATGCGCATGCGTGAAGGG 244
 DB 477 GAGAGAGACCTTGCGCGCTAGGCAAGACAGTGCAGCGCTCACCGGCTGGGTGCTTC 536
 QY 245 CAGCACTTTAAGTTATGACCAAAAGTCCCTGAGGCGCTTCGAGGTAAATGAAGCAT 304
 DB 537 GAGCACTTACCGGATATCAATTAATGAGTGAATTTAAGTCCCTTTGGCAAGATT 596
 QY 305 GGGCAGGATTTTAAATCATCAACGCGTACCGTTCTGCCACATGTGATCATCGGCTGG 364
 DB 597 ATCAACAGCTTTAAATCCTATGAGCAATGTTATGTTAGCTTTAATAGCAAAATGGG 656
 QY 365 AGACAGAGGAGACACAGTCCGCGAGAGGCGCAACTTCTGCTGGCGCACTCATGAGCA 424
 DB 657 AGATGCCAAAGACAAAGGTTGAGATGAGTCAAGCTCAAGCTCAATTAATGAATGAATCA 716
 QY 425 CAGAGTCTTCCGCGCGAGCGCTGATGACAGAGCTGGCCATAGCTGCTTCAAGCAA 484
 DB 717 AGTGG-----CACCACTTAATGATCAATTTGGAGCACTTGGCTTTTAAACCA 770
 QY 485 GAACGCAAGTGGCGGAGGAGTCTTTCAGACGATTTGAACGCTTTCATGAGTACGG 544
 DB 771 GAATTTCCATCTCGAAGAGCGGTGTCTGTCTTATTAAGAACTTAAACATTTTGG 830
 QY 545 CACCCAGACCTTAGTGTGGCGCTTAATPAACCAACAGTTGTGCACTTCTGGAGATCC 604
 DB 831 GGCTCAGCCACTAGTCACTAGCAAAATGATACACATTTGTGTCTGTTGGAGACTC 890
 QY 605 CACAGTTAATGTAGGAGAGCGCGCATCCAAACGCTGTGGAATCTACAAGCATGTAGG 664
 DB 891 CAACAGTCAGGTAGAGATCTGCAATATTGGCTATGATGAGATTTAAGACATGTGGG 950
 QY 665 GGATGATGCGCCGACAGACCTCCGTCGATGAGAGATGTTCTGCGC---TCGAATTTGG 720
 DB 951 AGAAAAAGTAGATGATCTTTAATAGAGAGAAATTCCTCCCTGATAGTAAAGAT 1010
 QY 721 CTATGTTGAGCAAAAGTTGACCAAGTCAACAGAGGCTTAATCTCACTTCAAGCC 780
 DB 1011 ATTGGCAAAATTTGATCAATGCAAAAGTTGAGGCGGTATATTGAGTCTTCAAGA 1070
 QY 781 TTAATAACAGAAATGGAATGAGTGGGCTTGAAGAGCGGCAATATTGGGTGAAGG 840
 DB 1071 TAAAGCTTGAATGATGAAGATCAATGATGAGAAATAGCCATCACTGATGATAGC 1130
 QY 841 AGCAGCCACAGAGATGATTAAGCGGCACTACCTCGCGCTTTCG-----887
 DB 1131 CTTCAGGTTCTGACCTTAATAACATCGGAATTCCTGCCAAGATGCAAGAAAGCTGG 1190

QY 888 -TCATCACTGCGCCCAAAACCAATGTGAACGATGTACCGGTGATCCGGCGCGGTAAAC 946
 DB 1191 TTCAGAGAGTGGCCCTTAAGTGTGAGAGTGTCTTCTTAAGAAAGAGGTGTGAGAGAGTTGA 1250
 QY 947 CATGGAATTTTTCGATCTAGCTTTAGTGTGTCCCGCAATGGAACATCTTCCACCGCTAA 1006
 DB 1251 TGAAGATGATTTTAAAGCTTTTACAGATGTCCCTTCTTATTCAGATTTATTTAGTGTG 1310
 QY 1007 GGACATGAGCAGATATCTACAGCAAGTATAGTATCATGATGATTAATAAGCAGAGCTG 1066
 DB 1311 AGAATCTGAAGAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1370
 QY 1067 GGAGAAACGCTGTGATGCTCTCAAGAAATCAAGGAGCTGCTCATTTTCAGCTATCAAC 1126
 DB 1371 GGATGAGCTGGCCAAATGCACTGAAGAAATTCGATCAGCTGTGT-----GCTGAGAC 1424
 QY 1127 TCAGCGCAGTTTGTGCTGTACAGCTAAAGAAATGTCTTAAGCTTGTGTGACATCTCT 1186
 DB 1425 TCCACAGTATGATGTCTTTTTCACATTTACGATTTGATGAGCACTTAACCTTTTC 1484
 QY 1187 CAAGGAGAACTACAGTACAGGTGATCCGCGAGGCGTGCATCACCATGCTACATGTC 1246
 DB 1485 AACTAAGATCTTATGATCCAGGTGTGAGAAAGCTTTATTAAGTGTAGCTGACCACTTTC 1544
 QY 1247 TAAAGCGCTGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1306
 DB 1545 AACAGTTTGGGAAACAGTTTGTATGATGATGCGGTGAAAGCATGTGATCACTTTTAA 1604
 QY 1307 TTTAATACGAACAGCGCGAGTCACTTGTGATGCTGCTTCCACATTAAGCTTCTGAATAT 1366
 DB 1605 TCTCGTCCCAATAGTGAAGAAAGTCAAGCACTTCTGATGTGACAGCATAGATTTAT 1664
 QY 1367 CATTAAGTATACAGTACAGCAAGTGTCAAGATCTACAGACACTGTGATCAATC 1426
 DB 1665 CATGCGCATCTACATGATCCAGACTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 1721
 QY 1427 AAAGTCGAAGACATTAAGTCCACACTGTGTGAGTGTGCTGCTTCCGAGAGTG 1486
 DB 1722 AAAATAGTCCCGTGAAGAGAGCTTCAATTTGAATTTTGAATTTATGTTGCAAGAGTG 1781
 QY 1487 GGAGAGAAAGCGGTGAAAGAAATGCCACCTGTAAGGAGCACTTAAATAATTCAT 1546
 DB 1782 GGAGACTCATTTATTTGAAGAAAGATGACCCGCTTGTGAAATTAATTAATTAATTAATTA 1841
 QY 1547 TGGGATGAGAGCTGGAGTGAACAGCGCCATTCAGATCGCTTATGAGCTTCCAGCGC 1606
 DB 1842 TATATGCTTGAAGCTGAGGCCAGAGTGAAGGCAAGAAAGCACTACATGGTCTTGAAG 1901
 QY 1607 TCACCTTCCAGAGCTGGCGATCAATATATATGAAACATTAGACATGAGCTGCCAGCGCGC 1666
 DB 1902 CCACCTTCCGTTGGAAGTGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1961
 QY 1667 ATTGAAGAGGAAAGAGAGGCGCGGAGAGAGAGAGAACTGTGATC 1711
 DB 1962 TCTTCAAACTTACTTAACAGATTTCTGCAGTGTGATGATCTTCC 2006

RESULT 6
 ADB61959
 ID ADB61959 standard; cDNA; 1908 BP.
 AC ADB61959;
 XX
 DT 04-DEC-2003 (first entry)
 DE Human cDNA encoding clone BRACE20079370.
 XX
 XX Human; ss; gene; pharmaceutical; diagnostic; gene therapy;
 KM tissue regeneration; cell regeneration; membrane protein;
 KM signal transduction-related protein; transcription-related protein;
 KM osteoporosis; neurological disease; cancer; tumour.
 XX
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH 158..1522
 FT CDS /tag=a
 FT /product="Clone BRACE20079370 protein"
 FT
 FT
 FN EPI308459-A2.
 XX
 PD 07-MAY-2003.
 XX
 PF 28-MAR-2002; 2002EP-00007401.
 XX
 PR 05-NOV-2001; 2001JP-00379298.
 PR 25-JAN-2002; 2002US-00350978.
 XX
 PA (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuho Y;
 XX
 DR MPI: 2003-450961/43.
 PS P-PSDB; ADB63929.
 XX
 PT New polynucleotides and polypeptides, useful for developing a diagnostic
 PT marker or medicines for regulation of their expression and activity, or
 PT as targets of gene therapy.
 XX
 PS Claim 1; Page: 222pp; English.
 XX
 CC The invention discloses a polynucleotide comprising a sequence selected
 CC from 1970 fully defined nucleotide sequences which encode novel
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
 CC or its partial peptide, an antibody binding to the polypeptide or peptide
 CC of the polynucleotide, immunologically assaying the polypeptide or
 CC peptide of the polynucleotide by contacting the polypeptide or peptide
 CC with the antibody of the encoded protein, and observing the binding
 CC between the two, a transformant carrying the polynucleotide in an
 CC expressible manner and an antisense polynucleotide. The oligonucleotide
 CC is useful as a primer for synthesizing the polynucleotide, or as a probe
 CC for detecting the polynucleotide. The polynucleotides and encoded
 CC proteins are useful as pharmaceutical agents and many disease-related
 CC genes may be included in them, for developing a diagnostic marker or
 CC medicines for regulation of their expression and activity, or as targets
 CC of gene therapy. The genes are involved in tissue and/or cell
 CC regeneration. Membrane proteins, signal transduction-related proteins,
 CC transcription-related proteins, disease-related proteins and genes
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
 CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
 CC the activity or expression of the encoded protein to treat diseases. The
 CC sequence presented is a cDNA of the invention. Note: Some of the sequence
 CC data for this patent is not represented in the printed specification, but
 CC is based on sequence information supplied by the European Patent Office.
 XX
 SQ Sequence 1908 BP; 531 A; 386 C; 411 G; 580 T; 0 U; 0 Other;
 XX

Query Match 2.3%; Score 116.6; DB 10; Length 1908;
 Best Local Similarity 48.7%; Pred. No. 3e-22;
 Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

QY 927 GGTGATGCGCGGCGCGTGAACATGATCTTTCAGATCTAGCTTGAAGTGTCCCGCA 986
 DB 353 GGAGGTGCTGGAGCGATGATGAATGATTTTAAAGCTTTTACAGATGTCCCTTCT 412
 QY 987 TTGAACATCTTCCAGCTTAAGACATGAGCATATCTCAAGCAGTACTAGTATCATC 1046
 DB 413 ATTGATTTTATTTAGTGAAGTCAAGAACTTAATTAATTAATTAATTAATTAATTAATTA 472
 QY 1047 AGTGATTAAGAGCAGACTGGAGAAACGTGTGATGCTCTCAAGAAATCAAGGCAATG 1106
 DB 473 TTAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 532


```

QY 1107 CTCATTCTGATATCAGCTGAGCGGAGTTTGTGCTGTAACGTAAGAATTGTCG 1166
DB 533 CTGTGTGCTGAG-----CTGCACAGTATGATGCTTTTTCACATTACGATGTG 586
QY 1167 TTAAGCTTGAGACATCTCAAGAGAGAACTACATCAAGGTATCCGAGGCGGAC 1226
DB 587 GATGAGACCTTAACCTTACAGTAAGATCTTAGATCCAGGTGTTAGAGAACTGTG 646
QY 1227 ATACCATGCGCTACATGCTCTAAGACGCTGAGAAATAAGTATGCTTGTGAGAC 1286
DB 647 ATTACTGTAGCCACCTTTCAACAGTTTGGAAACAGTTGATCATGAGGCGTGAACC 706
QY 1287 ATTTGGAGACCTGATTTATTTATATACGAAACAGCGGAGGATGATGATCCGCTTC 1346
DB 707 ATTGACCTACACTTTTATCTCGTCCCAATAGTGCAAAAGTATGCACTTCTGGA 766
QY 1347 ACAATAGCTGAGATATCATTAAGTATACATGACCAAACTGCTGCAAGTAC 1406
DB 767 TGTGACGACATGATTAATCAATGCGCATCTCAATGACCACTTATACCTTTAATA 826
QY 1407 ACAGACACTGATATCAATCAAGTGAAGACATAGGTCACACTGTGAGCTGATG 1466
DB 827 ACAACCATTTGCA-----CATCAATCAAGTCCCGTAGAGACCTTCAATTTGATTTTA 883
QY 1467 GTGCTGCTTTGAGAGAGTGAGACGAAAGCGTTGGAAGATGCCACTGATTAAG 1526
DB 884 GATTTATTTGTCAGAGAGTGAGACATCATTTGGAAGACATGACGCGCTTGTGTT 943
QY 1527 GACACCTTAAAAAATTCATTTGGCGATGAGACGCTGCGACGCGCGCATTTCCAGATAC 1586
DB 944 GAAATATTTAAAAAGGAATTCATATGATGCTGACGCTGAGCCAGATGAGGCAAG 1003
QY 1587 GCTTATTTGGGCTTTGAGGCGTCACTTTCCAGAGCTGCGGATCAATATATGAACATTA 1646
DB 1004 ACATCATGAGGCTTATGAACACATTTCTGTGTGAGCTGAACATTTATATTTCCCTT 1063
QY 1647 GACATAGCTGCGGAGCGGATTAAGAAAGGAGAGAGGCGGAGAGAGAGAACT 1706
DB 1064 GAGCCATCTTATCAAGAGAGTCTCAACTTATTAAGAGTTCTGGGAGTGTAGCATCT 1123
QY 1707 GGTAC 1711
DB 1124 CTTCC 1128

```

RESULT 7

ADB79874 standard; DNA; 5614 BP.

```

XX ADB79874;
XX AC
XX DT 04-DEC-2003 (first entry)
XX DE Rat putative KIAA0627 coding sequence. SEQ ID 114.
XX KM Analgesic; pain; streptozocin-induced diabetes; rat; gene; ds.
XX OS Rattus norvegicus.
XX FN EP1279744-A2.
XX PD 29-JAN-2003.
XX PF 26-JUL-2002; 2002EP-00255249.
XX PR 27-JUL-2001; 2001GB-00018354.
XX PR 07-FEB-2002; 2002GB-00002910.
XX PA (WARN ) WARNER LAMBERT CO.
XX PI Brooksbank RA, Dixon AK, Lee K, Plincock RD;
XX WPI; 2003-395407/38.

```

```

DR P-PSDB; ADB79873.
XX Use of isolated gene sequences and encoded polypeptides that are
XX upregulated in the spinal cord in response to streptozocin-induced
XX diabetes for screening compounds for the treatment of pain, or for
XX diagnosing pain.
XX Claim 1; Page 208-210; 334pp; English.
XX The present invention relates to nucleotide sequences which are useful in
XX the screening of compounds for the treatment of pain, or for the
XX diagnosis of pain. The nucleotide sequences are up-regulated in the
XX spinal cord in response to streptozocin-induced diabetes. The present
XX sequence is one such nucleotide sequence.
SQ Sequence 5614 BP; 1668 A; 1120 C; 1181 G; 1645 T; 0 U; 0 Other;

Query Match 2.3%; Score 116.6; DB 10; Length 5614;
Best Local Similarity 48.7%; Pred. No. 6.2e-22;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

QY 927 GGTATGCGCGCGCGGACATGAGATCTTTCGATCTGAGCTTGAAGTGTCCGCAA 986
DB 269 GGAAGTCTGAGAGCATGATGAAGATTTTATAAAGCTTTTACAGATGTCCCTCT 328
QY 987 TTGAACATCTTCAACGCTAAGACATGACGATATCTTCAAGCAAGTACTAGTATC 1046
DB 329 ATTCAATTTATTTCTAGTCAGAACTCGAAGAAATTAATAATCAGGAAATTTTG 388
QY 1047 AGTATTAAGCAAGCATGCGAGAAACGTGTGATGCTTCAAGAAATCAAGGCAATTG 1106
DB 389 TCAATATTAACATGATGCGGATTCAGCGGCAATGCAATGAAATTCATCACTG 448
QY 1107 CTCATTTCACTATCACTCAACCGCAGTTTGTGCTGTACAGCTTAAAGAAATTGTCG 1166
DB 449 CTGTGCTGTGAG-----CTGCACATGATGATGCTTTTTCACATTTACATTTG 502
QY 1167 TTAAGCTTGTGACATCTCAAGAGAACTACGATCAAGTATCCGAGCGCTGC 1226
DB 503 GATGAGACACTTAACTTCACTGATGATCTTAGATCCAGGAGTTAAGAAAGTGTG 562
QY 1227 ATGACCATGCGCTACATGCTTAAAGCGTGAGAAATTAAGTATGCTTGTGTGAGC 1286
DB 563 ATTACTGTAGCCCACTTTCAACAGTTTGGAAACAGTTTATGATGAGGCTGAAGCC 622
QY 1287 ATTTGAGACACCTGATTTAATTAATATACAGAACGCGGAGGCTTGCCTTCC 1346
DB 623 ATTGTACTACACTTTTATTTATCTGCTCCCAATATGCAAAAGTCACTTCTGGA 682
QY 1347 ACAATAGCTGGAATATCATTAAGTATACATGACACCAAGCTGCTCAAGATCTAC 1406
DB 683 TGTGACGACATCAAGATTTATCATTTGGCATATCATATACCAAGCTTATACCTTATA 742
QY 1407 ACAGACACTGTGAATCAATCAAGTGAAGACATTAAGTTCACACTGTGTAGCTATG 1466
DB 743 ACAAGCAATTTGCA-----CATMAAATCACTTCCGTGAGAGAGCTTCAATTTTGA 799
QY 1467 GTGCTGCTTTGAGAGAGTGAGAGCAAGAGCGTTGGAAGAGATGCCACTGTAAG 1526
DB 800 GATTTATTTGTCAGAGATGAGAGCTCATTTCTTGAAGAGATGAGCGCTTGTGTT 859
QY 1527 GACACCTTAAAAAATTCATTTGCGATGACGACTGCGATGACGCGCCCATTCAGATAC 1586
DB 860 GAAACTATTAAGGAATTCATGATGCTGACGCTGAGGCGCAGAGTGAAGCAAGAAAG 919
QY 1587 GCTTATTTGGCTTTGAGGCGTCACTTTCCAGACTGCGGATGAATATATATGAACATTA 1646
DB 920 ACATACATGGGCTTTGAAAACCACTTTCGTGAAAGCTGAAAACATTATTAATTCCTT 979
QY 1647 GACATAGCTGCCGAGCGCATTAAGAGGAAAGAGAGGCGGCGAGAGAGAGAACT 1706
DB 980 GAGCCATCTTATGAGAAAGTCTTCAAACTTAAAGTTCGCGAGTGTAGCATCT 1039

```


QY 1707 GGTAC 1711
 Db 1040 CTTCC 1044

RESULT 8
 ADX07257
 ID ADX07257 standard; DNA, 5614 BP.
 AC ADX07257;
 DT 21-APR-2005 (first entry)
 DE Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 1822.
 XX cytostatic; cyclin-dependent kinase; cdk; biomarker; gene; de.
 XX Homo sapiens.
 OS WO2005012875-A2.
 PN 10-FEB-2005.
 PD 29-JUL-2004; 2004WO-US024424.
 PF 29-JUL-2003; 2003US-0490890P.
 PR (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;
 PI WPI; 2005-163068/17.
 DR P-Psdb; ADX07258.
 PT Biomarkers useful for predicting or determining the response of a mammal
 PT to a cancer treatment comprising administration of a modulator of cyclin-
 PT dependent kinase activity.
 PS Claim 5; SEQ ID NO 1822; 141pp; English.

CC This invention describes a novel method of predicting or determining
 CC whether a mammal will respond or is responding to an anti-cancer agent
 CC that modulates cyclin-dependent kinase (cdk) activity. The method
 CC comprises measuring the level of one or more biomarkers selected from
 CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID
 CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the
 CC invention is utilized in a kit for determining or predicting whether
 CC patient would be susceptible or resistant to treatment by an agent
 CC modulating cdk activity. The invention also describes a method for
 CC utilizing individualized genetic profiles for treating diseases and
 CC disorders based on patient's response and molecular level, specialized
 CC microarrays comprising the biomarkers described, antibodies directed
 CC against the biomarkers and a cell culture model to identify biomarkers.
 CC The cdk modulator is preferably N-5-[(1,1-dimethylethyl)-2-oxazoly-1]methyl-2-thiazoly-4-piperidine carboxamide, 0.5-L-
 CC tartaric acid salt. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pat_sequences. This
 CC sequence encodes a biomarker used in the method of the invention.

CC Sequence 5614 BP; 1668 A; 1120 C; 1181 G; 1645 T; 0 U; 0 Other;
 SQ

Query Match 2.3%; Score 116.6; DB:14; Length 5614;
 Best Local Similarity 48.7%; Pred. No. 6.2e-22;
 Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

QY 927 GGTGATGCCGCGCCGCTTAACCATGAATCTTTGAACTAGACTTTGAGGTGTCGCGCA 986
 Db 269 GGAAGGTGCTGAGCACTTGTGAAGATGATTTATATAAGCTTTTACAGATGCTCTCT 328

QY 987 TTGAACATCTTCCACCGCTTAAGACATGACGATATCTACAGCAAGTACTAGTATC 1046
 Db 329 ATTCAAGATTATTTCTAGTCGAGAACTCGAAGAAACATTTAAATCAAGGAAATTTTG 388

QY 1047 AGTGATAAAAAGCAAGACTGGGAGAAACGTGTGATGCTTCAAGAAATCATGAGGCATTG 1106
 Db 389 TCAGATGATATAACATGACTGGATCAGCGTCCAAATGACATGAAGAAATTCATGACTCG 448

QY 1107 CTCATCTCAGCTATCAACACTCCAGCGGATTTGTGTGCTATACAGCTAAAGAAATTTGTCG 1166
 Db 449 CTGTGTGCTGAG-----CTGCACAGATGATATGCTTTTTCACATCTTACATGATTTTG 502

QY 1167 TTAAGCTTGTGGAATCTCTCAAGAGGAACTAGACATCAAGTATCCGAGGCGTGC 1226
 Db 503 GATGAGACATCTTAACCTTTCAGCTAAGATCTTGAATCCAGGTGTGTAAGAACTTGT 562

QY 1227 ATCAACATCGCTTACATGCTTAAGACGCTGAAGAAATTAACATGATGCTTCTGCGAGC 1286
 Db 563 ATTACTGTAGCCCACTTTCAACAGTTTGGGAAACAAAGTTGATCATGGGCTGTAAGCC 622

QY 1287 ATTTTGAAGACCTGATTAATTTAATACAGAACGCGGAAAGGTGATTCATCCGCTTCC 1346
 Db 623 ATGTACCTACACTTTTAAATCTGCTCCCAATGTGCAAAAGTATGCAACTTCTGGA 682

QY 1347 ACAATAGCTCTGAAGTATATCATTAAGTATACATGACACCAAGCTGCTCAAGATCTAC 1406
 Db 683 TGTGAGCAATCAAGTTTATCATTCGGCATATCATGATACCACTTATACCTTTAATA 742

QY 1407 ACAGACTCTGGAATCAATCAAGTCAAGAGACATAGCTCAACTGTGTAGCTGATG 1466
 Db 743 ACAAGCAATGTCA---CATCAAAATCAAGTTCCTCGTAGAGACGTTCAATTTGAATTTTA 799

QY 1467 GTGCTGCTCTTGAGAGTGGCAGACGAAGCGTTGAAAGAAATGCCAGCTACTAAGG 1526
 Db 800 GATTTATTTGTGCAAGATGGACAGCTATTTCTGAAAGACATGACAGCCCTCTTGTT 859

QY 1527 GACACCTTAAAAAATCCATTTGGCGATGACAGATGCGATGACGCGCCATTCAGATAC 1586
 Db 860 GAACATTTAAAAAGGAAATTCATGATCTGACGCTGAGGCCAGAGTGGAGCAAGAAAG 919

QY 1587 GCTTATTTGGCTTTGAGCGCTGACCTTTCAAGACTGCGGATCAAAATATATGAACATTA 1646
 Db 920 ACATCATGAGGTCTTAGAAGAACCACTTCTCTGTGAAGCTGAAGAACTTTATATATTCCTT 979

QY 1647 GACATAGTGGCCCGCGGCACTTGAAGGGAAGGAGGCGGCGGAGAGAGAACT 1706
 Db 980 GAGCCATCTTTATCAGAAAGATCTTCAAACTTATTAAGATTTCTGGCAGTGTAGCATCT 1039

QY 1707 GGTAC 1711
 Db 1040 CTTCC 1044

RESULT 9
 AA158443
 ID AA158443 standard; cDNA, 6445 BP.
 XX AA158443;
 AC
 XX
 DT 22-OCT-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 646.
 XX
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 XX peripheral nervous system; neuropathy; central nervous system; CNS;
 XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 XX leukaemia; ss.
 XX Homo sapiens.
 OS WO200153312-A1.
 PN 26-JUL-2001.

CC defined (51) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041,
 CC 1372, 3996, 3945, 1798, 585, 1782, 927, 5714 or 2282 nucleotides as
 CC given in the specification, its translated or protein coding portion, its
 CC extracellular portion or its active domain. The GPCR-like polypeptides
 CC and polynucleotides are useful for the treatment of diseases of
 CC ophthalmic, neurological, immunological and nephritic systems. They may
 CC also be used to treat hormonal dysfunction, cancer, atherosclerosis and
 CC diabetes. The antibodies are useful for detecting or quantitating the
 CC polypeptide in tissue. The polypeptides can also be used as molecular
 CC weight markers and as a food supplement. This sequence represents a human
 CC polynucleotide of the invention.

XX Sequence 6445 BP; 1955 A; 1241 C; 1315 G; 1934 T; 0 U; 0 Other;

Query Match 2.3%; Score 116.6; DB 5; Length 6445;

Best Local Similarity 48.7%; Pred. No. 6.8e-22; Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

Qy 927 GGTGATGCCGGCCGCTAACCATGGAATCTTTCGATCTAGCTTTGAGTGTCCCGCA 986
 Db 362 GGAGTGCTGGAGCAGTTGATGATGATTTTATAAAGCTTTACAGATGTCCTTCT 421
 Qy 987 TTGAACATCTCCACGCTAAGACATGACATATCTACAGCAAGTACTAGTATCATC 1046
 Db 422 ATTCAGATTTTATCTAGTCGAGAACTCGAAGAACTTAAATCAAGGAAATTTTG 481
 Qy 1047 AGTGATTAACCGCAGACTGGGAGAAAGCTGTGATGCTCTCAAGATCAGGCATTTG 1106
 Db 482 TCAGATGATTAACATGATCGGATGATGATGATGATGATGATGATGATGATGATG 541
 Qy 1107 CTCATCTCAGCTATCAGACTCAGCCGAGATTTGCTGCTACAGCTTAAAGAAATTTG 1166
 Db 542 CTGTGCTGGAG-----CTGCACAGATATGATCTTTTCAACATTTACGATTTTG 595
 Qy 1167 TTAAGCTTCGTGACATCTCTCAGAGGACATACATACAGTATCCGAGGCGTGC 1226
 Db 596 GATGAGACATTAACCTTTCAGCTTAAGATCTTAATCCAGGTGTTAGAGAACTTGT 655
 Qy 1227 ATTCACATCGCTACATGCTTACAGCGCTGAGAAATACTAGTCCCTTCGTGAGAG 1286
 Db 656 ATTTAGTGAAGCCACCTTTCAACAGTTTGGAAACAAAGTTGATCAGCGCTGAAGC 715
 Qy 1287 ATTTGAGCACTGATTAATTAATTAACAGAAACAGCGGAGAGTTCATCGCTTCC 1346
 Db 716 ATTTGACCTACACTTTTATATCTGCTCCCAATAGTCGAAAGATGCAACTTCTGGA 775
 Qy 1347 ACAATAGCTGGAATATATCTTAAGTATACATGCAACCAAGCTGCTCAAGTTCAC 1406
 Db 776 TGTGACGACATCAGATTTATCATCTGGAATCTCATGTAACCAAGCTTAATACCTTTAATA 835
 Qy 1407 ACAGACATCTGGAATCATCAATCAAGTGAAGACATAGAGTCCACACTGTGAGCTGANG 1466
 Db 836 ACAAGCAATGCA---CATCAAAATCAGTCCCTGAGAGAGACCTTCATTTGAATTTTA 892
 Qy 1467 GTGCTGCTCTTCGAGAGTGGAGAGAGAGCGGTGGAAGAGATGCCACTTAAGG 1526
 Db 893 GATTATTTGTTCCAGAGATGGAGAGATCTATTCATTTGAAAGACATGACCGCTTGGTT 952
 Qy 1527 GACACCTTAAAAAAATTCATTTGGAGTGAAGCTGCGATGACGCGCCCATTCAGATAC 1586
 Db 953 GAAACTATTAATAAAGGAATTCATGATGCTGAGCGTGAAGCCAGATGGAAGCAAGAG 1012
 Qy 1587 GCGTATTTGGGCTTTAGGCGGCTTCACAGAGCTGGCGGATCAATATATGGAACATTA 1646
 Db 1013 ACATACATGAGGCTTAGAAACCATCTTCTGCTGAGCTGAAGCTTAATATATTCCTTT 1072
 Qy 1647 GACATAGCTGCCACCGCATTAAGAAAGGAAACGAGAGGCGCGGAGAGAGAACT 1706
 Db 1073 GAGCATCTTATCAAGAGTTCGAACTTAATTAAGAGTTCTGGAGGTGATCATCT 1132
 Qy 1707 GGTAC 1711
 Db 1133 CTTC 1137

RESULT 11
 ID ADB48414
 ID ADB48414 standard; cDNA; 6445 BP.

AC ADB48414;
 DT 04-DEC-2003 (first entry)
 DE Novel human cDNA SEQ ID NO 324.

ss; cancer; neurodegenerative disease; human.

Homo sapiens.

US2003104529-A1.

05-JUN-2003.

04-JAN-2002; 2002US-00037270.

21-JAN-2000; 2000US-00488725.

25-APR-2000; 2000US-00523117.

19-JUL-2000; 2000US-00620312.

(ZHOU/) ZHOU P.

(TANG/) TANG Y T.

(LIU/) LIU C.

(ASUNDI/) ASUNDI V.

(DRMANAC/) DRMANAC R T.

Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;

WPI; 2003-678194/64.

New polynucleotide, useful for treating diseases e.g., cancer or

neurodegenerative diseases.

Claim 1; SEQ ID NO 324; 99p; English.

The invention relates to a polynucleotide comprising a sequence given in

the specification, or its mature protein-coding portion, or its

complement. The polynucleotide is useful for treating diseases e.g.,

cancer or neurodegenerative diseases and many others listed in the

specification. The present sequence represents a novel human cDNA. Note:

The sequence data for this patent did not form part of the printed

specification but was obtained in electronic format directly from USPTO

at seqdata.uspto.gov/sequence.html?docID=20030104529.

Sequence 6445 BP; 1955 A; 1241 C; 1315 G; 1934 T; 0 U; 0 Other;

Query Match 2.3%; Score 116.6; DB 9; Length 6445;

Best Local Similarity 48.7%; Pred. No. 6.8e-22; Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

Qy 927 GGTATGCCGGCCGCTAACCATGGAATCTTTCGATCTAGCTTTGAGTGTCCCGCA 986
 Db 362 GGAGTGCTGGAGCAGTTGATGATGATTTTATAAAGCTTTTACAGATGTCCTTCT 421
 Qy 987 TTGAACATCTCCACGCTAAGACATGACATATCTACAGCAAGTACTAGTATCATC 1046
 Db 422 ATTCAGATTTTATCTAGTCGAGAACTCGAAGAACTTAAATTAAGGAAATTTTG 481
 Qy 1047 AGTGATTAACCGCAGACTGGGAGAAAGTGTGATGCTCTCAAGATCAGGCATTTG 1106
 Db 482 TCAGATGATTAACATGATCGGATGATGATGATGATGATGATGATGATGATGATG 541
 Qy 1107 CTCATCTCAGCTATCAGCTACCGCAGTGTGCTGCTACAGCTTAAAGAAATTTTG 1166
 Db 542 CTGTGCTGGAG-----CTGCACAGATATGATGCTTTTCAACATTTACGATTTTG 595
 Qy 1167 TTAAGCTTCGTGACATCTCTCAGAGGAACTACATCAGAGTATCCGAGGCGTGC 1226

```

Db      596 GATGAGCACTTAACTTTCAGCTAAGATCTTAAATCCAGTGTAGAGAACTTGT 655
Qy      1227 ATCACCATGCGCTACATGCTTAAAGACGCTGAGAAATAAATCTAGTCCCTTCTGTGAGC 1286
Db      656 ATTACTGTAGCCCACTTTCACAGTTTGGGAAACAGTTGATCATGCGCCCTGAAGC 715
Qy      1287 ATTTGGAGCAGCTGATTAATTAATATACGAAACAGCGCGAAGTCAATGATCCGCTTC 1346
Db      746 ATTGTACCTACACTTTTATCTCGTCCCAATAGTCAAAAGTCAATGCAACTTCTGGA 775
Qy      1347 ACAATAGCTCTGAATATCTATCTATAGTATACATGACCCAACTGCTCAAGATCTAC 1406
Db      776 TGTGAGCAATGATTAATTAATCTAGTGGCACTCACTGATGACCACTTAATACCTTAATA 835
Qy      1407 ACAGACACTCTGATCAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1466
Db      836 ACAGCAATGCA---CATCAAAATCAGTTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 892
Qy      1467 GTGCTGCTCTTGGAGAGTGGCAGACGAAAGCGTTGGAAAGATGCCAGCTAAAG 1526
Db      893 GATTATTTGTTGCAAGAGTGGCAGACTCATTCATTGAAAGACATGACAGCCGCTTGGTT 952
Qy      1527 GACACTTAAATAAATCCATTTGCGATGAGACTGCGATGACAGCCGCCATTCAGATAC 1586
Db      953 GAAACTATTAAATAAGGAAATTCATGATGCTGACGCTGAGGCCAGAGTGGAGCAAGAAAG 1012
Qy      1587 GCGCTATTTGGGCTTTCAGGCGCTCACTTTCAGAGCTGGCGGATCAATATATGAAACATTA 1646
Db      1013 ACATACATGAGGCTCTTGAAGAACCACTTCTCTGAGAGCTGAAACATTAATATTCCTT 1072
Qy      1647 GACATAGCTGCGCCAGCGGCACTTAAGAAAGGAAAGAGGCGGAGAGAGAGAGAACT 1706
Db      1073 GAGCATCTTATACAGAGAGTCTCAAACTTAAAGAGTTCTGGAGAGTATGACATCT 1132
Qy      1707 GGTAC 1711
Db      1133 CTTC 1137

```

RESULT 12

AA158442 ID AA158442 standard; cDNA, 6487 BP.

XX AC AA158442;

XX DT 22-OCT-2001 (first entry)

XX DB Human polynucleotide SEQ ID NO 645.

XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000MO-US034263.

XX PR 23-DEC-1999; 99US-00471275.

XX PR 21-JAN-2000; 2000US-0048725.

XX PR 25-APR-2000; 2000US-00552317.

XX PR 20-JUN-2000; 2000US-00598042.

XX PR 19-JUL-2000; 2000US-00620312.

XX PR 03-AUG-2000; 2000US-00653450.

XX PR 14-SEP-2000; 2000US-00662191.

XX PR 19-OCT-2000; 2000US-00693036.

```

PR      29-NOV-2000; 2000US-00727344.
XX      (HYSE-) HYSEQ INC.
PA      Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX      Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI      Zhou P, Goodrich R, Dymnack RT;
XX      WPI; 2001-442253/47.
DR      P-PSDB; AAM39286.
XX      Novel nucleic acids and polypeptides, useful for treating disorders such
PT      as central nervous system injuries.
PS      Claim 1; SEQ ID NO 645; 10078BP; English.
CC      The invention relates to human nucleic acids (AA157798-AA161369) and the
CC      encoded polypeptides (AAM38642-AAM42213) with noctropic.
CC      immunosuppressant and cytostatic activity. The polynucleotides are useful
CC      in gene therapy. A composition containing a polypeptide or polynucleotide
CC      of the invention may be used to treat diseases of the peripheral nervous
CC      system, such as peripheral nervous injuries, peripheral neuropathy and
CC      localised neuropathies and central nervous system diseases, such as
CC      Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC      lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC      utilisation of the activities such as: Immune system suppression,
CC      Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC      and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC      assays for receptor activity, arthritis and inflammation, leukemias and
CC      C.N.S disorders. Note: The sequence data for this patent did not form
CC      part of the printed specification
XX      Sequence 6487 BP; 1969 A; 1248 C; 1327 G; 1943 T; 0 U; 0 Other;
SQ
Qy      927 GGTGATCCGCGCGGCTAACGATGAAATCTTGAATCTGAGTGTGAGTCCCGCA 986
Db      362 GGAGCTGCTGAGAGAGTGTAGTGAAGATGATTTTAAAGCTTTTACAGATGTCCTTCT 421
Qy      987 TTGAACATCTTCCAGCGTAAAGACATGAGATCTCAAGACAGTACTGATCATC 1046
Db      422 ATTCAGATTTATTTCTAGTCAAGATCTGAAAGAACTTAATTAATAACAGGAAATTTTG 481
Qy      1047 AGTATTAATAACGAGAGTGGAGAAAGTGTGATGCTCTCAAGAGATCAAGGAGATTG 1106
Db      482 TCAGATTAATAACGAGAGTGGAGAAAGTGTGATGCTCTCAAGAGATCAAGGAGATTG 541
Qy      1107 CTCATTTCAAGCTATCACTCAAGCGGAGTGTGCTGATCAGTAAAGAAATTTG 1166
Db      542 CTGTTGCTGAGAG-----CTGACAGATGATGCTTTTCAACATTTACGATTTGTTG 595
Qy      1167 TTAAGCTTGTGAGATCTTCAAGAGAGAACTAGATCAAGAGTATCCGAGAGCGTGC 1226
Db      596 GATGAGCACTTAACTTCAAGCTTAAGATCTTAAGATCCAGGAGTGAAGAGCTTGT 655
Qy      1227 ATCAGATGCGCTACATGCTTAAAGACGCTGAGAAATAAATCTAGTCCCTTCTGTGAGC 1286
Db      656 ATTACTGTAGCCCACTTTCACAGTTTGGGAAACAGTTGATCATGCGCCCTGAAGC 715
Qy      1287 ATTTGGAGCAGCTGATTAATTAATATACGAAACAGCGCGAAGTCAATGATCCGCTTC 1346
Db      746 ATTGTACCTACACTTTTATCTCGTCCCAATAGTCAAAAGTCAATGCAACTTCTGGA 775
Qy      1347 ACAATAGCTCTGAATATCTATCTATAGTATACATGACCCAACTGCTCAAGATCTAC 1406
Db      776 TGTGAGCAATGATTAATTAATCTAGTGGCACTCACTGATGACCACTTAATACCTTAATA 835
Qy      1407 ACAGACACTCTGATCAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1466
Db      836 ACAGCAATGCA---CATCAAAATCAGTTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 892

```

Oy	1467	GTGCGCTCTTCGAGGAGTGGCAGACGAAGGCGTTGGAAAGAAATGCAACCGTACTAAG	1526
Dd	893	GATTATTGTTGCGAAGAGTGGCAGACTATTTCATTGGAAAGACATGACGCCGTCTTGTT	952
Oy	1527	GACACTTTAAAAAAATTCATTGGCGCATGCGATGCGACGCCGCAATTCAGATAC	1586
Dd	953	GAACCTATTAAAAAGGAATTCATGATCTGACGCTGAGGCCAGAGTGGAGCAAGAAAG	1012
Oy	1587	GCCATTGGGCTTTAGAGCGTCACCTTTCCAGAGCTGGCGGATCAAATTATGGAACATA	1706
Dd	1013	ACATACATAGGGGTCTTAGAAACCACTTCTCGTGAAGCTGAAACATTTATTAATTCCTT	1072
Oy	1647	GACATAGCTGCCCGACGCGGCATTAGAAAGGAAGACAGAGGCGCGGAGAGAGAACT	1706
Dd	1073	GAGCCATCTTATATCAGAAAGTCTTAAACTTAAAGAGTTCTGSCAGTGTAGACTCT	1132
Oy	1707	GGTAC	1711
Dd	1133	CTTCC	1137

RESULT 13

ADQ98653	ADQ98653	standard; cDNA; 6487 BP.
ADQ98653	ADQ98653	
AC	ADQ98653	
XX		
DT	23-SEP-2004	(first entry)
XX		
De	DNA encoding human GPCR-like protein seqid 323.	
XX		
Kw	ophthalmological; immunomodulatory; cytosolic; antiatherosclerotic;	
Kw	antidiabetic; GPCR-like protein; ophthalmic disorder;	
Kw	neurological disorder; immunological disorder; nephritic disorder;	
Kw	hormonal dysfunction; cancer; atherosclerosis; diabetes;	
Kw	molecular weight marker; food supplement; human; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	US6569662-B1.	
XX		
PD	27-MAY-2003.	
XX		
PE	19-JUL-2000; 2000US-00620312.	
PR	21-JAN-2000; 2000US-00488725.	
PR	25-APR-2000; 2000US-00552317.	
XX		
PA	(HYSE-) HYSEQ INC.	
Xx		
PI	Tang YT, Zhou P, Drmanac RT;	
XX		
DR	WPI; 2001-442255/47.	
XX		
PT	New G-protein-coupled receptor-like polypeptides and polynucleotides,	
PT	useful for treating diseases of ophthalmic, neurological, immunological	
PT	and nephritic systems and hormonal dysfunction, cancer, atherosclerosis	
PT	and diabetes.	
PS		
PS	Example 2; SEQ ID NO 323; 92pp; English.	
XX		
CC	The invention describes an isolated polynucleotide (i) comprising a fully	
CC	defined (s1) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041,	
CC	1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as	
CC	given in the specification, its translated or protein coding portion, its	
CC	extracellular portion or its active domain. The GPCR-like polypeptides	
CC	and polynucleotides are useful for the treatment of diseases of	
CC	ophthalmic, neurological, immunological and nephritic systems. They may	
CC	also be used to treat hormonal dysfunction, cancer, atherosclerosis and	
CC	diabetes. The antibodies are useful for detecting or quantitating the	
CC	polypeptides in tissue. The polypeptides can also be used as molecular	
CC	weight markers and as a food supplement. This sequence represents a human	
CC		

CC	polynucleotide of the invention.
XX	
SO	Sequence 6487 BE; 1969 A; 1248 C; 1327 G; 1943 T; 0 U; 0 Other;
	Query Match 2.3%; Score 116.6; DB 5; Length 6487;
	Best Local Similarity 48.7%; Pred. No. 6.8e-22;
	Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;
OY	927 GGTATATCCGGCGCCGTAAACCATGGATCTTTTGGAACTTAGCTTTGAGTGGTCCGCCAA 986
DB	362 GGAGTGCTCGAGCGATGTGATGAAGATGATTTTATAAAGCTTTTACAGATGTCCCTTCT 421
OY	987 TTGAACATCTTCCACGCTAAGACATGAGCGATATCTTACAAGCAAGTACTAGTATCATC 1046
DB	422 ATTACAGTTTATTTCTAGTCCAGAACTCGAAGAAACATTAATATAATACAGAAATTTTG 481
OY	1047 AGTATATAAAACGACACTGGAGAAACGTGTGATGCTCTCAAGAAATCAGGCAATG 1106
DB	482 TCAGATGATTAACATGACTGGATCAGCGTCCCAATGCACTGAAGAAATTCGATCACTG 541
OY	1107 CTCAATTTCTAGCTATCACTCACTCAGCCGCACTTGTGCTGTACAGCTAAAGAAATTCG 1166
DB	542 CTGTGTTCTGAG-----CTGCAAGTATGATTTGCTTTTCAACATTAACGATTTGTG 595
OY	1167 TTAAAGCTTGAGCATCTCCCAAGGAGAACTAGATCAGAGGTATCCGCGAGCGCTGC 1226
DB	596 GATGAGACACTTAAACTCTTACGCTAAGATCTTATGATCCAGGTGTGTTAGAAAGCTTGT 655
OY	1227 ATCAACATCCGCTTACATGTCTTAAAGCCGTGAGAAATTAATAGATGCTTGTCTGAGCC 1286
DB	656 ATTACTGTACCCACCTTTCAACAGTTTGGAGAAACAAAGTTGATCATGGCGCTGAAGCC 715
OY	1287 ATTTTGGAGACCTGATTAATTATTAATACAGAACAGCGCGAAGTCAATGCTCCGCTTCC 1346
DB	716 ATTGTACCTTACAGTTTAACTTCGTCTCCCAATAGTCAAAGTCAAGGCAACTTCTGGA 775
OY	1347 ACAATTACTCTGAAGTATATCATTAAGTATACATGSCACCAAGCTGCTCAAGATCTAC 1406
DB	776 TGTGACGCAATCAGATTTATCATTTGCGCATATCTCATGTACCCAGACTTAATACCTTTA 835
OY	1407 ACAGACACTCTGAATCATCAAAAGTCAGAGAACATTAAGTCCACACTGTGTAGCTGATG 1466
DB	836 ACAAGCAATTGCA---CATCAAAATACGTTCCCGTGAGAGACGTTCAATTTGAATTTTA 892
OY	1467 GTGCTGCTCTTTCAGAGAGTGGCAGACGAGGCGTTGGAAAGAAATGCCACCGTACTAAG 1526
DB	893 GATTATATGTTGCAAGAGTGGCAGACTCATTTCAATTGAAAGACATGACGCGTCTTGATT 952
OY	1527 GACACCTTAAAAAATCATTTGGCGATGACAGCTGCGATGACACGCGCCCATTCACAGTAC 1586
DB	953 GAAACTTATTAATAAAGGAATTCATGATGCTGACGCTAGGCGACAGTGGAGGCAAGAAAG 1012
OY	1587 GCCATATGGGCTTTACAGCGTCACTTTCAGAGCTGCGGATCCAAATATATGAAACATTA 1646
DB	1013 ACATATCATGGTCTTAAAGAACCACTTTCTGTGTAGAGCTGAACATATATATATTCCTT 1072
OY	1647 GACATAGCTGCCACCGCGCATTAAGAAAGGAACGAGAGGCGCGGAGAGAGAGAACT 1706
DB	1073 GAGCACTTATATGAGAAAGATGCTTCAAACTTAAAGAGTTCTGGCAGGTGATGACTCT 1132
OY	1707 GGTAAC 1711
DB	1133 CTTCCT 1137
RESULT 14	
ID	ADB48413 standard; cDNA; 6487 BP.
XX	ADB48413;
XX	ADB48413;
DT	04-DEC-2003 (first entry)
XX	

DE Novel human cDNA SEQ ID NO 323.
 XX ss; cancer; neurodegenerative disease; human.
 KW Homo sapiens.
 OS US2003104529-A1.
 XX
 XX
 XX PD 05-JUN-2003.
 XX PF 04-JAN-2002; 2002US-00037270.
 XX PR 21-JAN-2000; 2000US-00488725.
 XX PR 25-APR-2000; 2000US-00552317.
 XX PR 19-JUL-2000; 2000US-00620312.
 XX
 PA (ZHOU/) ZHOU P.
 PA (TANG/) TANG Y T.
 PA (LIUC/) LIU C.
 PA (ASUNDI/) ASUNDI V.
 PA (DRMANAC/) DRMANAC R T.
 XX
 FI Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;
 DR WPI; 2003-678194/64.
 XX
 PT New polynucleotide, useful for treating diseases e.g., cancer or
 PT neurodegenerative diseases.
 XX
 PS Claim 1; SEQ ID NO 323; 99pp; English.
 XX
 CC The invention relates to a polynucleotide comprising a sequence given in
 CC the specification, or its mature protein-coding portion, or its
 CC complement. The polynucleotide is useful for treating diseases e.g.,
 CC cancer or neurodegenerative diseases and many others listed in the
 CC specification. The present sequence represents a novel human cDNA. Note:
 CC the sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html?docid=2003104529.
 XX
 SQ Sequence 6487 BP; 1969 A; 1248 C; 1327 G; 1943 T; 0 U; 0 Other;
 Query Match 2.3%; Score 116.6; DB 9; Length 6487;
 Best Local Similarity 48.7%; Pred. No. 6,8e-22;
 Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

QY 1347 ACAATAGCTGGAAGTATATCATTAAGTATACATGCAACCAAGCTGCTCAAGATCTAC 1406
 DB 776 TGTGACAGCAATTCAGATTTATCATTCGGCATCTCATGTATACCAAGATTATACCTTTAATA 835
 QY 1407 ACAGACACTGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1466
 DB 836 ACAAGCAATTCGA---CATCAAAATTCAGTTCCCGTGAAGAGACCTTCATTTGAATTTTGA 892
 QY 1467 GTGCTGCTCTTCGAGAGATGGCAGACAGAGCGCTTGGAAAAGAAATGCAACCGTACTAAGG 1526
 DB 893 GATTATTTGTTGCAAGATGGCAGACATCTCATTTGTAAGAAAGATGACAGCCGCTTGTT 952
 QY 1527 GACACCTTTAAAAAATTCATTTGCGATGACAGACTGCGATGCAAGCGCCGCTTCAGATAC 1586
 DB 953 GAACTATTTAAAAAGGGAATTCATGATGCTGACCGCTGAGCCAGAGTGGAGCAAGAAAG 1012
 QY 1587 GCCTATTGGGCTTTCAGGCGCTCACTTCCAGAGCTGGCGGATCAATATATGAAACATTA 1646
 DB 1013 ACATACATGGGCTCTTAAAGAACACACTTCTGTTGTAAGCTGAACATTTATATATTCCTT 1072
 QY 1647 GACATAGTGGCCAGCGCGCATTTAGAAAGGAAAGGAGCGCGGAGAGAGAGAACT 1706
 DB 1073 GAGCCATCTTTATTCAGAAAGAGCTTCAAACTTAAAGAGTTCTGGCAGGTATGACATCT 1132
 QY 1707 GGTAC 1711
 DB 1133 CTTC 1137

RESULT 15
 AA160228
 ID AA160228 standard; cDNA; 6491 BP.
 XX
 XX AC AA160228;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 4217.
 XX
 KW Human; nootropic; immunosuppressant; cytostratic; gene therapy; cancer;
 KW peripheral nervous system neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSF-) HYSF INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
 PI Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 DR WPI; 2001-442253/47.
 DR P-PSDB; AA41072.

XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 4217; 10078bp; English.

XX The invention relates to human nucleic acids (AA157799-AA161369) and the
CC encoded polypeptides (AA157799-AA161369) with neurotrophic,
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX

Sequence 6491 BP; 1965 A; 1251 C; 1328 G; 1947 T; 0 U; 0 Other;

Query Match 2.3%; Score 116.6; DB 4; Length 6491;

Best Local Similarity 48.7%; Pred. No. 6.8e-22;

Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

Qy 927 GGTGATGCCGGCGCCGTACCAATGGAATCTTTCGATCTAGCTTTGAGTGTGCTCCGCA 986
Db 362 GGAGGTGCTGGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 421
Qy 987 TTGAACATCTTCCAGCTAAGACATGAGCATGATCTCAAGACATGATCTAGTATCATC 1046
Db 422 ATTGATGATTTATTCTAGTCGAGACTCGAAGAACATTAATTAATCAGGAAATTTTG 481
Qy 1047 AGTGATTAACCGCAGACTGGAGAAACGTGTGATGCTCTCAGAGATCAGGCGATG 1106
Db 482 TCAGATGATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 541
Qy 1107 CTCATTTGAGCTATCACTGACGCGCATTTGCTGCTGATGATGATGATGATGATGATG 1166
Db 542 CTGTGCTGAGG-----CTGCACAGATGATGATGATGATGATGATGATGATGATGATG 595
Qy 1167 TTAAGCTTGTGAGATCTCTCAAGAGAACTAGATCAGATGATGATGATGATGATGATG 1226
Db 596 GATGAGACATTAACCTTTCAGTAAAGATCTTAAATCCAGGTGTGAGAGAACTGT 655
Qy 1227 ATCAACATGCTGCTATGATGCTTAAGAGCGTGAAGAAATAAATGATGCTTGTGAGAG 1286
Db 656 ATTACTGATGCCACCTTTCAGACATTTTGGAAACAGTTGATCATGGCGCTGAAGC 715
Qy 1287 ATTTTGAGCAGCTGATTAATTAATTAACAGAACAGCGGAGAGTTCATGCTCCGCTTC 1346
Db 716 ATTTGACCTACCTTTTAATCTCTCCCATAGTGCAGAAAGTTCATGCAACTTCTGGA 775
Qy 1347 ACAATAGCTGAAATATATATCTTAAGTATACATGACCAAGCTGCTCAAGATCAC 1406
Db 776 TGTGAGCAATGATTAATTAATCTGAGCACTCAATGACCAAGCTTAATACCTTAATA 835
Qy 1407 ACAGACATGCTGATCAATCAAGTGAAGACATAGCTCCACATGCTGAGCTGATG 1466
Db 836 ACAACCATTTGCA---CATCAAAATCAGTCCCGTGAAGAGACGTTCAATTTGAATTTTA 892
Qy 1467 GTGCTGCTTGTGAGAGTGCAGACGAAGCGTTGGAAGAGATGCCACGCTACTAAG 1526
Db 893 GATTATATGTTGCAAGAGTGCAGACATCTATGGAAGACATGACGCTCTTGTT 952
Qy 1527 GACAGCTTAATAAATCAATGCGATGAGACGCTGATGACGCGGCAATTCAGATAC 1586
Db 953 GAAACTATTAATAAAGGAATTCATGATGCTGACGCTGAGGCCAGATGAGGCAAGAAAG 1012
Qy 1587 GCGTATGCGGCTTTCAGCGCTCACTTTCAGAGCTGCGGATCAATATATGAAACATTA 1646

Db 1013 ACATACATGGGTCTTAAACCACTTCTGTGTAAGCTGAAACATTAATTCCTT 1072
Qy 1647 GACATAGTGGCCGAGCGCGCATTPAGAAAGGAACGAGAGGCGCGGAGAGAGAACT 1706
Db 1073 GAGCATCTTATCGAAGAGTCTTCAAACTTACTTAAGAGTTCTGCGAGTGAATCT 1132
Qy 1707 GGTAC 1711
Db 1133 CTTC 1137

Search completed: November 29, 2005, 15:58:23
Job time : 1891 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OW protein - protein search, using sw model

Run on: November 23, 2005, 15:02:54 (Search time 61 seconds
(without alignments)
2022.166 Million cell updates/sec

Title: US-10-030-850-2

Sequence: 1 MAYRRPSDLDGFIQMPKAD.....RNCISGGSSSTKNSSAASSS 1492

Scoring table:

Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA:
1: /cgn2_6/ptodata/1/1aa/5-COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/6-COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/H-COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/BCRUS-COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/RE-COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1236.5	16.2	255	US-09-270-767-32225	Sequence 32225, A
2	1236.5	16.2	255	US-09-270-767-47442	Sequence 47442, A
3	1214.5	16.0	916	US-10-104-047-2297	Sequence 2297, Ap
4	920	12.1	723	US-10-104-047-2487	Sequence 2487, Ap
5	503.5	6.6	454	US-10-104-047-2083	Sequence 2083, Ap
6	248	3.3	51	US-09-270-767-61363	Sequence 61363, A
7	248	3.3	66	US-09-270-767-45835	Sequence 45835, A
8	208.5	2.7	1025	US-09-248-796A-20573	Sequence 20573, A
9	196.5	2.6	2283	US-10-172-502-4	Sequence 4, Appl1
10	193.5	2.5	2137	US-09-134-001C-4463	Sequence 4463, Ap
11	191	2.4	4377	US-09-949-016-6978	Sequence 6978, Ap
12	182	2.4	3913	US-09-949-016-10933	Sequence 10933, A
13	173.5	2.3	2781	US-09-698-295-10	Sequence 10, Appl1
14	173.5	2.3	2907	US-09-698-295-11	Sequence 11, Appl1
15	172	2.3	2733	US-09-949-016-11433	Sequence 11433, A
16	172	2.3	3259	US-09-949-016-6507	Sequence 6507, A
17	171.5	2.3	1140	US-09-538-092-647	Sequence 647, Appl
18	171	2.2	907	US-08-938-830-26	Sequence 26, Appl
19	171	2.2	907	US-09-020-222-26	Sequence 26, Appl
20	171	2.2	907	US-09-068-377-26	Sequence 26, Appl
21	167	2.2	1237	US-09-862-027-78	Sequence 78, Appl
22	163.5	2.1	787	US-10-104-047-3340	Sequence 3340, Ap
23	162.5	2.1	686	US-09-252-991A-20509	Sequence 20509, A
24	162.5	2.1	2842	US-07-741-940-7	Sequence 7, Appl1
25	162.5	2.1	2842	US-08-289-548A-7	Sequence 7, Appl1
26	162.5	2.1	2842	US-08-452-654-7	Sequence 7, Appl1
27	162.5	2.1	2842	US-08-449-731-7	Sequence 7, Appl1

ALIGNMENTS

28	162.5	2.1	2843	1	US-08-452-655B-2	Sequence 2, Appl1
29	162.5	2.1	2843	1	US-08-452-655B-7	Sequence 7, Appl1
30	162.5	2.1	2843	2	US-08-450-582-2	Sequence 2, Appl1
31	162.5	2.1	2843	2	US-08-450-582-7	Sequence 7, Appl1
32	162.5	2.1	2843	2	US-09-538-092-1007	Sequence 1007, Ap
33	162.5	2.1	2973	1	US-08-821-355A-7	Sequence 7, Appl1
34	162.5	2.1	2973	1	US-09-003-687A-7	Sequence 7, Appl1
35	162.5	2.1	2973	2	US-08-136-605-7	Sequence 7, Appl1
36	161.5	2.1	313	2	US-09-252-991A-29545	Sequence 29545, A
37	161.5	2.1	2843	2	US-10-092-138A-30	Sequence 30, Appl1
38	161.5	2.1	2843	2	US-08-681-219A-30	Sequence 30, Appl1
39	161.5	2.1	3696	2	US-09-134-001C-5080	Sequence 5080, Ap
40	161	2.1	2663	2	US-09-538-092-1252	Sequence 1252, Ap
41	160.5	2.1	526	2	US-08-252-991A-28761	Sequence 28761, A
42	158	2.1	1401	2	US-09-976-594-1035	Sequence 1035, Ap
43	157.5	2.1	564	2	US-09-949-016-6898	Sequence 6898, Ap
44	157.5	2.1	564	2	US-10-069-540A-2	Sequence 2, Appl1
45	157	2.1	565	2	US-09-949-016-6902	Sequence 6902, Ap

RESULT 1
US-09-270-767-32225
Sequence 32225, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 255
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 32225
LENGTH: 255
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-32225

Query Match 16.2% Score 1236.5; DB 2; Length 255;

Best Local Similarity 96.5%; Pred. No. 1.1e-95; Indels 1; Gaps 1;

Matches 247; Conservative 0; Mismatches 8;

QY 859 LISLTQYADGKELTQOLKCVLDMFRKMFNDTHKYSLEFLDTVELILVHNETSRNG 918

DB 1 LISLTQYADGKELTQOLKCVLDMFRKMFNDTHKYSLEFLDTVELILVHNETSRNG 59

QY 919 SSSCITRLFNKLGTDLNSMHSKIWTLOVHYEFPLOLQKELFRITSDSTQTPTTKR 978

DB 60 FLITLRLFNKLGTDLNSMHSKIWTLOVHYEFPLOLQKELFRITSDSTQTPTTKR 119

QY 919 IAILRFLDLANTCKSSDPDSOACERTYLKLAADOKSMELRSQASCLVALYN 1038

DB 120 IAILRFLDLANTCKSSDPDSOACERTYLKLAADOKSMELRSQASCLVALYN 179

QY 1039 INTPOMTLLILDLPVYODSARSCISHMRROSGCNSGANSPPSSSPKPLQSPS 1098

DB 180 INTPOMTLLILDLPVYODSARSCISHMRROSGCNSGANSPPSSSPKPLQSPS 239

QY 1099 VGPFASLQSHHQSLI 1114

DB 240 VGPFASLQSHHQSLI 255

RESULT 2
US-09-270-767-47442

Sequence 47442, Application US/09270767

Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 47442
LENGTH: 255
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-47442

Query Match 16.2% Score 1236.5; DB 2; Length 255;
Best Local Similarity 96.5% Pred. No. 1.1e-95;
Matches 247; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 859 LISTOYADGKELTQOOLKCVLDFRMPDHTKTVSLFDYTELILVHNETSNG 918
DB 1 LISTOYADGKELTQOOLKCVLDFRMPDHTKTVSLFDYTELILVHNETSNG 59
QY 919 SSSCITLFLNKLGTDLNLSMHSKIWKTLQVHVEYPTLOLKEFLRIISDSTQPTTKTR 978
DB 60 LFLITLRLFNKLGTDLNLSMHSKIWKTLQVHVEYPTLOLKEFLRIISDSTQPTTKTR 119
QY 979 IAIIRFLTDLANTYCKSSDFPSDOSQACERIVLKLQALADQKSMELRSQARSCVALYN 1038
DB 120 IAIIRFLTDLANTYCKSSDFPSDOSQACERIVLKLQALADQKSMELRSQARSCVALYN 179
QY 1039 LNTQOMTLLADLPVYODSARSCHSHMRROSOSCNCSGSPSSPLSSSPKPLQSPS 1098
DB 180 LNTQOMTLLADLPVYODSARSCHSHMRROSOSCNCSGSPSSPLSSSPKPLQSPS 239
QY 1099 VGPFASLQSHHQLSI 1114
DB 240 VGPFASLQSHHQLSI 255
RESULT 3
US-10-104-047-2297
Sequence 2297, Application US/10104047
Patent No. 6943241
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. 6943241e1 full length cDNA
FILE REFERENCE: HI-A0105
CURRENT APPLICATION NUMBER: US/10/104,047
CURRENT FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER:
PRIOR FILING DATE:
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2297
LENGTH: 916
TYPE: PRT
ORGANISM: Homo sapiens
US-10-104-047-2297

Query Match 16.0% Score 1214.5; DB 2; Length 916;
Best Local Similarity 32.8% Pred. No. 7.3e-93;
Matches 321; Conservative 180; Mismatches 322; Indels 157; Gaps 29;

QY 9 LDGFIQMPKADMRVKVQLAEDLVTLSDDTNSIVC---TDMGFLIDGL-MPWLTGSHF 63
DB 5 MESCLAQVLQKDVGRKLVQVQELIDYFSDKOKSADLEHDMLEKLVGLATSWNSNY 64
QY 64 KIAQKSLFAFSELKRLGSDNATYATLPHVIRIGDSRTPVPEKQULLRDMEHVYL 123
DB 65 KVVLLGMDILSALVRLQDRFKAQIGIVLPSLIDRLGSAKOSVREDDQTLTKLMD-OAA 123
QY 124 PQAALIDPLATSCFKNKAKVREEFLLQIVNALHEVGTQOLSVRYIIPVQALLGDPVTN 183
DB 124 NPQVYVMDMLOG-FRHNKFRREGICLCIATLANSGQTLTSLKIVYHICNLGDPDSQ 182
QY 184 VREAIQTLVEIYKVGDRLRDLRRMDVPASKLAMLEQKFDQVQKQGLLPSALKNTN 243

DB 183 VRDAIINSLVEIYRHVGERVRADLSK-KGLPQSRLANVIFTKFDEVOKSGNMVIOQA----- 236
QY 244 GNGVGLDEADNIGLRSEPTMIRKPLHSAVSSSLRPKPNVDVYG----- 288
DB 237 -NDKNFDESDVD-GNRPS-----SASTSKAPPSRRVNGMOTTRRLSGSTLQSK 286
QY 289 -----DAGAVTWSEPFSSFEVVPOLNIFRAKMDDIYKQVLIISDKNADWEKRVADLK 342
DB 287 SSAKEGAGAVDEEDFIKAFDDVPVQIYSSRLDEESINKIREILSDKHDEQVNAALK 346
QY 343 KIRALLLSYHTQOPAVOLKELSLFVDILKEELRSQVIREACTITAMSKTLRKLD 402
DB 347 KIRALLLAGAAYDNFQO-HLRLLDGA-FKLAKDLRSQVREACTITLGHLSVLAGKFD 404
QY 403 AFGSILEHINLIQNSAKVIAASTALKYIIKYTHAPKLKIYDTLQNSKSDIRST 462
DB 405 HGAELIPTFNLIPNSAKIMATSGVAVVAILIRHTIIPRLIPITNSCT-SKSVAVRR 463
QY 463 LCEIMVLLFEEWQTKALERNATVLRDTLKKSIGDADCDARRSHRYAWAFRRHPELADQ 522
DB 464 CFEFLDLLQEWQTHSLERHISVLAETIKGIDHDESEARIEAKCYWGFHSFBSREAEH 523
QY 523 IYGTLDIAQALEREKREGGGGGGTGTGTATETRTVGRIGTPTLOKPTSMRSIS 582
DB 524 LYHTLESYKALQSHLK-----NSDSIVSLPQSDRSS--SSQSLMRPLSAKISPT 574
QY 583 AVPTAAQRAKVAQVLYRSORRKLPGPNNSNOASMTGAAGSLPRPRINSNGTPT-- 640
DB 575 GSTTSRASTVSTSVSTTGLOR-----SRDIDVNAALSA--SKYSSSGTTPFS 624
QY 641 ---ATPGRSV-----TPPRGRAG---V 657
DB 625 SAALPPGSYASLESRRHREDMEYIGHDSGRIRTRQSSGATVASTPNNRGSRRAKV 684
QY 658 SGOPOGSRSTSPSTKLRDQVGI-GNYTRG-ATGAIIRKXSGIPRSTASRETSPTSGG 715
DB 685 SOSOPGSRSSPGKLSGGYGLTGGSSRPVTPSSSEKSKLPRSGCCRETSPPNIG- 743
QY 716 GLMKRSWYSTGAGS-RTTFRNPNVPRPSAPARLLAQSRGAHTLGVDDQDPDYVSGDYM 774
DB 744 -----LDRFGLOPGHIPPSSVANM-----KVLSTDELAHV-----ADAL 779
QY 775 RSGGMRGRKLMGRDESDIDSEASVCSERSFDSYTRGNKSNYSLSGHTRLDWSTQR 834
DB 780 EK-PYRRRYEPYGVSDDDANSASSVCSERSYGS--RNGIPIHY-----L 822
QY 835 APFDITITIOFCASTWSEKRGGLISLTOYLDGKELTQOOLKCVLDFRMPDHTK 894
DB 823 RQTEDEVAEVLNHCASSNWSERKGLLQWLKSORLTSVELKRLCEIFTRMADPHSK 882
QY 895 -VYSLFDYTELILVHANE 913
DB 883 RVFSMLETIVDFTIIHKD 902

RESULT 4
US-10-104-047-2487
Sequence 2487, Application US/10104047
Patent No. 6943241
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. 6943241e1 full length cDNA
FILE REFERENCE: HI-A0105
CURRENT APPLICATION NUMBER: US/10/104,047
CURRENT FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER:
PRIOR FILING DATE:
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2487
LENGTH: 723
TYPE: PRT

ORGANISM: Homo sapiens
US-10-104-047-2487

12.1%; Score 920; DB 2; Length 723;

Query Match 32.0%; Pred. No. 3.3e-68; Mismatches 243; Indels 98; Gaps 19;

Best Local Similarity 144; Mismatches 274; Indels 98; Gaps 19;

Matches 243; Conservative 144; Mismatches 274; Indels 98; Gaps 19;

5 KPSDLGFIQOMPKADRVKVLAEPLVFL-SPDTNSIVCTDMGFL---IDGLMPWLTG 60
2 EPMSEMFCAQOVQKQVGRLOVQGGELLVLGAGALSDEEDLGRKGTVDALTVGVGS 61
61 SHEKIAQKSLAEFSEIKRLGSDFNATYATVLPVHIDRLDSDQTYREKQQLIRLMEH 120
62 SNRVSLIMGLLEIISAFDRILSTRFKSYVAVYVALIDRMGDADKXDEAQTILKLM-D 120
121 RVLPQALLDKATSCFKHNAKAREEFLDTIVNALHEGYTQOVSVVYTPPCALLGDP 180
121 QVAPPMYIWEQLA-SGFHKHNFERSREGVCLLETLIFGAQPLVLSKLIPIHLCILFGDS 179
181 TVAVREAAIOTLVEIYHGVGDRLEPDLRRMDVPAKTLAMEQKFDQVQKQGLLPSALK 240
180 NSQVRDAIILAIYEIYHVEKRMVLYR-GIPPALEMIIPAKFDEVOSSGGMILSVCK 238
241 NTNGNGVGLDEADNIGLRERT-----RMIRPLHSVSSSLRKPENVNVT-- 287
239 DKS-----FDDESSVD-GNRPSSAASAFKVPAPKTSGNPANSARAKPGSAGGPKVGASKE 292
288 GDGAVTMSFESEFEVVPQNTIFHAKMDIYKQVVIISDKNADMEKVDALKIRAL 347
293 GGAGAVDEDEFIKAFVTVPISQIYSRELETLNKIRELISDDGHWDQANALKKIRSL 352
348 LI---LSYHTOPQFVAVQVKELSLFVDLKEELRSQVIREACTIIYMSKTLKNDLAF 404
353 LVGAQAQYDFQHLRLDGLAKLS-----AKDLSQVVEACTIVAHLSVLGNKRDHG 407
405 CWSLLEHLNLIONSAXVIASTIALKYIITKTHAPRLKIYDTLNQSKSKDIRSTLC 464
408 AEAIVPLFNLVPASAKVATSGCAIRFIIRHVRPLIPLINSCT-SKSVVRRRSF 466
465 ELAVLLEFEMOTKALENNATYLRDLTKSIGDACDARRHRSRYVMAFRHFPFLADQY 524
467 EFLDLLQEWOTHSLEHRAALVETIKGIHDADEARVAKTYMGLRNHFPEAETLY 526
525 GTLDIAAQRALEREREGG-----GGTGTGTGTAPETRR 559
527 NSLEPSYQKSLQTYLNSGVSASLPQSDRSSSSQESLNRFSSKMWSTANSTVAGRVA 586
560 TVSRIGTPTGTLQKPTPSMRISAVDTAAQRAVRAQYTLYSQKPELGNNSNQAS-- 617
587 GSSKASSLPGLQ-----RSRSDIDVNAAGAKAHHAAGSVRSR-LGAGALNAGSYA 639
618 -----MTGAAAGSLPRRLNSNGCTATPGSTPRPRAG---VSOSOPGSR 666
640 SLEDTSKLDGTASDEGRVRAKLSA-----PLAGKNAKADSRSRSTXWVSQSPGSR 694
667 TSPSTKLRDQYGGIYNYRATGAI PKKASGIPRSTAS 705
695 GSPGRVL-----TTTALSTVSSGVQRYLVNS 720

RESULT 5
US-10-104-047-2083
Sequence 2083, Application US/10104047
Patent No. 6943241
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. 6943241el full length cDNA
FILE REFERENCE: H1-A0105
CURRENT FILING DATE: 2002-03-25
CURRENT APPLICATION NUMBER: US/10/104,047
PRIOR FILING DATE:
PRIOR APPLICATION NUMBER:
NUMBER OF SEQ ID NOS: 4096

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2083
LENGTH: 454
TYPE: PRT
ORGANISM: Homo sapiens
US-10-104-047-2083

Query Match 6.6%; Score 502.5; DB 2; Length 454;
Best Local Similarity 32.5%; Pred. No. 2.4e-33; Mismatches 138; Indels 54; Gaps 12;

Matches 138; Conservative 77; Mismatches 156; Indels 54; Gaps 12;

261 PTRMIRPLHSVSSSLRKPENVNVTGAGAVTMSFESEFEVVPQNTIFHAKMDIY 320
44 PANSARKP-----GSAGGRVAGASNEGAGVDEDFIKATFDVPSIQYSRELETL 98
321 KQVLIISDKNADMEKXVDALKIRALLI---LSYHTOPQFVAVQVKELSLSVDLKEE 377
99 NKIREILSDQKHDMDQANALKKIRSLVGAQAQYCCFQHLRLDGLAKLS-----AKD 153
378 LRSQVIREACTIIYMSKTLRNLDAFCWSILEHLNLIONSAXVIASTIALKYIIRK 437
154 LRSQVIREACTIIYMSKTLRNLDAFCWSILEHLNLIONSAXVIASTIALKYIIRK 437
438 THAPKLLIYDTLNQSKSKDIRSTLCMLVLLFEEMOTKALENNATYLRDLTKSIGDA 497
214 THVFRILPLITSCT-SKSVPRRRSFEDLILQEMQHSLEHRAALVETIKGIHDA 272
498 DCDARRHRSRYVMAFRHFPFLADQYGTLDIAAQRALEREREGGCGTGTAPET 557
273 DAAARVAKTYMGLNHFPEAETLYNSLEPSYQKSLQTYLKS--SVASLPQS 336
558 RRTVSRIGTPTGTLQKPTPSMRISAVDTAAQRAVRAQYTLYSQKPELGNNSNQAS 617
327 DRESS---SOSLELNRFSSKMWSTANSTVA---GRVSA-----GSSKAS 365
618 MTGAAAGSLPRR-----LNSNGCTPATPGSVTRPRGRAGVSOSOPGSRSTPSYTL 673
366 ---SLPGSIQKRSRSDIDVNAAGAKAHHAAGSV--RSRLAAGALNAGSYA---SLGK 416
674 RDQYG 678
417 RTDYG 421

RESULT 6
US-09-270-767-61363
Sequence 61363, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 61363
LENGTH: 51
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-61363

Query Match 3.3%; Score 248; DB 2; Length 51;
Best Local Similarity 98.0%; Pred. No. 1.5e-13; Mismatches 49; Indels 1; Gaps 0;

Matches 49; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

697 GIPSTASSRETSPTRSGGGLMKRSMYSTGASRRTERPNNVPASAPAR 746
1. GIPSTASSRETSPTRSGGGLMKRSMYSTGASRRTERPNNVPASAPAR 50

RESULT 7
US-09-270-767-45835

```

; Sequence 45835, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 45835
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-45835

```

```

Query Match          3.3%; Score 248; DB 2; Length 66;
Best Local Similarity 98.0%; Pred. No. 2.4e-13;
Matches 49; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Cy 697 GIPSTASRETSPTRSGGLMKRSMYSTGASRRTPRRNNPVPSAPAR 746
Db 1 GIPSTASRETSPTRSGGLMKRSMYSTGASRRTPRRNNPVPSAPAR 50

```

```

RESULT 8
US-09-248-796A-20573
; Sequence 20573, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20573
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (9), (11)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno
US-09-248-796A-20573

```

```

Query Match          2.7%; Score 208.5; DB 2; Length 1025;
Best Local Similarity 20.1%; Pred. No. 6.1e-08;
Matches 250; Conservative 208; Mismatches 429; Indels 357; Gaps 64;

```

```

Cy 329 DKNADMEKRDVALKKIRALLILSYTOPOFAVOLKELSLFVDILKEELSOVIREACI 388
Db 38 ETEBNMKREKNILOMRKILNGSNATOPRSELYOCITITANGMCKGASSLTITLSSNCQ 97
Cy 389 TIAYMSKTLRNKLDAFCMSILEHLINTL--IONSAXVIASATIALKYIYTHAPKLLK 445
Db 98 LKSCAVILKKSLEFVASLEFPTLIKCSSTKNIINSTANMSVALVANYLPYT--SKMIQ 155
Cy 446 IYTLTNGSKSDIRS-TLCELMVLLFE-----EMQTKALERNAVTVAROLKLSIG 495
Db 156 RIT-LASDRNYQOPSSYLILWLHLLIKIGIDRSYIGHDSSFLEANKVMKLLK---- 210
Cy 496 DADCDARRSHRYAVYAFRRHFPELADQIYGTLDINAORALR-EREKGGGGGTGTGTGTA 554
Db 211 DANENVQTAKECYWCFTFRVPEDEARLLKLEPNIVYALERSQRESG-----GSGIA 263

```

```

Cy 555 PETRRVTSRIGRTPTGCT-PSMKISAVDTAAQRAVRAQYTLXSRQKLGPNNS 613
Db 264 P--IRTLSS-----SRPSRSLKE-----ALLENKX-----ELRRRPPRSNG 299
Cy 614 NQASMTGAASGSLPRPLRNSNSGCPATTPGVTTPRPGAGVS-OSOPSRSTS---P 669
Db 300 EQST--KIKSVPLRP-----TKSSRLKSLRPDVGHKSQPAVRASWTP 345
Cy 670 STKLRDQYGIIGNYRGATGAIPIKASGIPRSTASRETSPTSGGLMKRSMYSTGAGS 729
Db 346 STQ-----SG-PKATPKQERSKTE----- 364
Cy 730 RRTPERNNPVPSAPAPALAQREAEHTLVGDGDPDYVS--GDYMRSGMMGKRM 786
Db 365 ---VHKSPLEISPPSRL--DTGAVSSFNKKDPMINFLSSSDSLIKEINILKYAI 419
Cy 787 GRDESDDIDSEASVCSERFSDSYTGNKSNYSLSGSH-----TRLDMSTQRAPF 837
Db 420 GK--ENLPSILNSLK-----SISKHYQFMKPLTSDNYTKKALL 460
Cy 838 ---DIIETIIQPCASTWSEKDGILSLTOYLDGKELTQOOLKCVLDMFKMMDTHTK 894
Db 461 LRPD---FLRVCAIV-FDEPDEAVISLI-----ICCI-----DVLT- 493
Cy 895 VSIPLDTVELLIVANETSNGSSCLTRLEN-KLG-TDLNMSMSKIYKTIQVNAHE 952
Db 494 ---FYESACNLTLVADTPNIGSHALVMOISNOKLITKLI-----IQALLIALSKH 543
Cy 953 PPLOLKELEPRIISDSQTPTTKTRIAILRPLTLANTYCKSDSPSDSQACERTVLK 1012
Db 544 AVTDVQGEIYFQEI-----VKLIIVKATDIYSLIC----- 574
Cy 1013 LAQLAADQKMELESRQASCLVALYNLTPQMTLLADLPKRYQVDSARSC--IHSNMRQ 1070
Db 575 -----QLFRQ-----LYTIDSNKFLSLVDVGGKLEVEVFIYGIESTVLE 616
Cy 1071 SGCNSGANSPPSSPSSSPKPL-QSPSVQPPFASLOSHHOLSTSPSRSSSVQOE 1129
Db 617 RPSSKLDYDVLVVKFTNLIGTFVLSQKPSADDFMLPKRSEFLGSGEYVNSKXVAKIEA 676
Cy 1130 LTPS---SELDIOHNOKTSEERHCFGGQOYOTALAPNGNGLVYHDOGOQDSCASLS 1185
Db 677 LDSPQKQSQMD-QISLKRNSB-----HQBSEVLVDPE-ANVSATDGG----- 717
Cy 1186 SNSKTQSANTTOSNTPEASATMLDNL-----BERTONAKSPSTD----- 1227
Db 718 -GSKLNGSKN--DDDNLKQEMERIDPLKPISNKIRKISYEDAKQNEKPLERNMGQFQ 774
Cy 1228 ---AKYITVSINNAENGELI-----LASNLMESEVVRVALITLKQPPVELLOTS--- 1273
Db 775 YAKFSRAIRVNA-VAENMSLTSQEFESCCSKLVETPQSALILKVT--WFDVSLSVSSYDY 831
Cy 1274 ---LTNIGICIKGNCGLPKHFRSIRMLNLLEAEHTDVVAGLHVLSKIMSNNKMRH 1330
Db 832 QEEFLNG-----KHLKESLWEPFSKIDKLDHS-LVMNGFLKLOLKFN----- 876
Cy 1331 NMHFELELLIKLIQCYQSHKEALRDISMIPIRAPSPLDLS-----INIVNPIV 1381
Db 877 DTPRVVDLFPALVDTGQ-----BELDSELYFIWNNMLSLNDELMSPEFKVALN--Y 928
Cy 1382 ATGEFFPNLCALIKILE-----VTEHHGSEITDAH-LDIYFPNLABADQSGVYKRAAV 1435
Db 929 LBEG-ENNLTLISCLNLYLAKVYVDDCDLVAKIYRLDTFGKLFH-----EREVFRKSAT 984
Cy 1436 FCYIKLYFVGEKVKPKPL-SVLPSPK-----VRLNLYYIEK 1471
Db 985 ICSNLT--LANTVSEVVDITDKVSKRYPATYGRILIEFTMR 1025

```

```

RESULT 9
US-10-172-502-4
; Sequence 4, Application US/10172502
; Patent No. 6841154

```



```
Qy      995 SDFPESDGOACERTYTLKLAQLAADOKSMELRBOASCLVALYNLNTPOWTLILLADLPKY 1054
Db      1755 VSDSNRASTSLSDSTSVSDSTSASTSBSASTSBE-----SESTASSTSSESTSTS 1808
Qy      1055 YODSARSCHTSHMRROSQSCNGCANGPSSSPPLSSSPKPLOSPGVPPFASJGHHHQLSI 1114
Db      1809 VSDSTSTSTSDASASTSTSESDBNSEBTSLESTSTSVSDSTSASTASASTSTSVSDSNS 1868
Qy      1115 SSTT-PRSRQSSVEOELLFSSELIDICHNIQKT-SEIRHCFGGOYTALAPNGFNHLOY 1172
Db      1869 ASTSLSGSTSTSVSDSTSTSASTSTSESDSDASTSLSGSTSTISDS-----1920
Qy      1173 HDQGQODSCASLSNKNTOSSANTTOS-----NTPSAMRLNDLERETTONAKS 1223
Db      1921 -----TSTSTSDASASTSTSESASTSTSVSESDSESTSVSESSSTSVSDSTSTSESAST 1975
Qy      1224 PTDAKAVITVSINMAANGELILASNMESVVAVALTLTFDOVV----ELLQTSLNMLGI 1279
Db      1976 STSESESTSESTSVSESSSTSTSIDSSSTSTSMSTSETFTSQSFINESOPTIGDSLSEDPT 2035
Qy      1280 CIKGNCELPRNK 1291
Db      2036 VTQSXTNMTNLNK 2047

RESULT 11
US-09-949-016-6978
; Sequence 6978, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6978
; LENGTH: 4377
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6978

Query Match          2.5%, Score 191, DB 2, Length 4377;
Best Local Similarity 18.1%; Pred. No. 2.3e-05;
Matches 303, Conservative 242, Mismatches 613, Indels 518, Gaps 76;

Qy      69 SLEARSELIKRGSPFNATATYLVLRVIDRLDGRDPTVEKKQOLLRLDMEHHRVLDPOL 128
Db      906 SLGARASASLRGSSD-RSITT-----LNRSTARSMMWIE-BLLVPSEKQHULTFRREF 966
Qy      129 IDKL-----ATSCFKHN---AKVREEPIQT-IVNAL--HEVYTOOLSVRVYIP----- 172
Db      957 SDLSLHYSWAADPTLDNVNLVSPHSIHGFLVSPFVDADRGSMDRSRHGMRIIIPRKCTA 1016
Qy      173 ----VCALI-----GDPTVNNREAAIQTLVELYKHVGDRLRDLARMDDVP----- 214
Db      1017 PTRITCRVLKRHKLANPPhGERGGISRLVENGPAGQFLGVL---VEIPHGSMRGK 1073
Qy      215 ASKLAML-----EOKFPDVKQEGLLPPLSKVTNGNGVGLDEADNIGLREERPTRMI 265
Db      1074 ERELVLKRENENETWEHQFD-S-KNEDL-----TELLNGMDELBDPEELG-KKRICRII 1126
Qy      266 KR-PLHSAVSSSLRRKPN-VNDVTGACAVTMSESSF-----EVVPO 307
```

Dh	1127	IKDPFOYFAVSRIRKOSNNOIGPBGGLISSTYPLVQASPEGALTRIKVGLQAQV	1185
Qy	308	LNIFHAKMDDIYQVL-----YIISDKADWEKRY-----	338
Dh	1186	-----DEIVKIIIGNKATFSPVITVEPRRRKFKPIITMTIPVPPSGEGVNGYKG	1236
Qy	339	DALKIRALLISHTQOPAVQKLKS-----LSFY-----DIL	374
Dh	1237	DTENLRILCSITGTSF-----AQMEDITGTPPLTFIKDCVSTTVNSABFWLADCHQVL	1292
Qy	375	KE-ELRSOVINEACITTAIYMSK-----TLNKIKDAFCWS--LIEHILNIIQNSAKV	422
Dh	1293	ETVGLAOLVIRE-LICPYNAKFVVPKAMDPEVSSLRCCMDDKDKLXLEQENEEV	1351
Qy	423	IASASTIAIK-----YIIKYTHAPRLK-----IYDTLNO-----SKSADRSTLC	466
Dh	1352	ARSDIEVLBESKPIYVDOCYGNLAPLTKGGQOLVENFSPKEMRLPFSIKIRDTQSOBCGR	1411
Qy	467	MVLLPEEMQYALERNATV-LRDTL-----KKSIGDADCDARHNR--YAYMARRRPPE	518
Dh	1412	LSPFKEPKRTKGLPQTAVCMINILTPAKKKEITESDODEIEKTRDROSFASTLARKYSY	1471
Qy	519	LADQIYGLDIAAQRALEREREGGGGGGTGCTGTAETRRVTRIGRTGTLCKPPSM	578
Dh	1472	LTF-----PQMIKRSIGAT	1485
Qy	579	RSISAVDTAAQAKVRAQYTLYSRQKRLGPNNNSQA-----SMTGAASGSLPPRRL	632
Dh	1486	RSLP-----TYS--YKPFSTRPYQSWTAPITVPQAKSGF---TSL	1524
Qy	633	NSNNGGTPATP-----GSVTPRP-----KGRAGVSGSQPSGRSTSEBTKLRDQYGGI---	680
Dh	1525	SSSSSNTPSASPLKSIWSVSTPSIKSTLSTASITTSYKSIISDVASPIRSRTWSSPIKTV	1584
Qy	681	-----GKYR--GATGRIPKKASGIPRSTASRETSRTRGGGMMKSMYSTG	726
Dh	1585	VSGSPNIOVSSGTLAPAVTEKTPK--GLASNSFFSRTSIVYTAGSLERSSIT--	1640
Qy	727	AGSRRTERRNNPVPSAPARLLAQSRERAEHTLGVGDGDQDPDYSGDYMSGMMGRKLM	786
Dh	1641	-----MTP-----PASPKSNIMYSSSLPFK-----SIITSAAPLISBPLK	1676
Qy	787	G-----RDESDDDISEASVCSSEBSPSSSTRGNKSNYSLSGSHTRLDWSTQAPFDIE	841
Dh	1677	SVSPVSRVADVISSAKITWASSLSSEPVKOMPGHAEVALVNGSISPLKYASS-----S	1729
Qy	842	TIIOFCASHTMSEKODLISLYOVLADGKELTLQOQKLCVLDMPKMMMDHTKXYSFLD	901
Dh	1730	TLINGCKAT-----ATLOEKISATNSVSSVSAATDYVEKVFSTTTAMPFSLRS	1780
Qy	902	TVTELLIVHANETSRNGSSCLTFLPFKLGTDLINSMSHKIMKTLQVHVEFFTOLOKE	961
Dh	1781	YVS--AAPSAFQISRTSASA--LYTSLOSS--ISATTSVTSIIIVPYSVNVLPBR	1834
Qy	962	LFRIIISDTQTPPTKTIRIARLFTDLA-----NTYKSSDFPSDOSQACERTVYLKLAQL	1016
Dh	1835	ALKTLDPDNS--FTKSAALALSPIKTLTTETHPOPHFSRSTSPKVSLSFLAPSAKLSTP	1892
Qy	1017	AADQSMELNSQASC-----LVALNLMTPQMTLLADLP-----KYU	1055
Dh	1893	SSLSOSSOILKDVAKEMEDIMRTAILIQTIVPEKKPQPELPEKGRIDDEBPFIKEVK	1952
Qy	1056	QDSAR-----SCISHMRQOSQNSGANSPPSSPLSSSPKPLQSPSPVGFASLOS	1107
Dh	1953	EDLYKVEILKXDVNCKGSPKSPKDKG--HSEDDMWIEFSSGEIRKA-----	2000
Qy	1108	HHHOLISSTSPRSRQSSVEOELLFSSLDIOHNIQKTSSEIRHCGGQYQIALAPNGFN	1167
Dh	2001	--RQAAASOSP--SLPERVQVKKAAASBKD--YMLTVIDYLTNDIGSSSLT-----	2047
Qy	1168	GHLOY-HDQGOQDSBALSNSKTQSSANTQSTNPSATMRLDNIRE-----	1215
Dh	2048	NLYKXEDAKQOEGQKVLPRALLOEHKLMR--PAMRRSTSEKELCMAADSFTGT	2105

```

QY 1216 -----RTTQAKSPDADAKVITVS---INNAENGELLASINMESEVVALTLT 1262
DB 2106 DTILESPPDFSGHDDKSPPLSDSGFETSEKTPSPAPOSAETTPGPKPLFHEVPIPVITET 2165
QY 1263 KQAPPELLQTSLTNIGICIKGNCELPN-----KHFPS 1295
DB 2166 KREVVHVRISYDPSAG-----DVPQTQPEEPVSPKSPFTFMELEPPTTSIKKVK 2217
QY 1296 IMRMNLINLEAHTVIVAGLHVLSKIMSNGMRHNMHFLILIKTIQCYOHSKEALR 1355
DB 2218 AFQOMASSEDH-----RVLSKGMVKEETH-----ITTTBMVYHSPGGE 2261
QY 1356 DIDSMPRIAPSLPLDLININVPVATGEPTNLCAITILEVTEHHGSEITDHLDIY 1415
DB 2262 GASERIE-----TMSVHDIMKAFOSGRDPS-----KELAGLEPHKSA-----V 2300
QY 1416 PNLARSADDDTQSMRKAAVFCIVLVLEGEKVKPKLTVANPSKVRLLNYIEK 1471
DB 2301 SPDVHKSAAETSAQHAER-----DNQMKPKLE-----RIIEVHIEK 2336

RESULT 12
US-09-949-016-10933
; Sequence 10933, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10933
; LENGTH: 3913
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10933

Query Match 2.4%; Score 182; DB 2; Length 3913;
Beet Local Similarity 18.1%; Pred. No. 0.00011;
Matches 304; Conservative 241; Mismatches 613; Indels 518; Gaps 77;

```

```

DB 722 -----DEIVKILGNKATFSPVIVTEBRKKFHPIITMILVPEPPSGEVSNGYKG 772
QY 339 DALKIRALLILSHTOQPOFAVOLKELS-----ISFV-----DIL 374
DB 773 DTTNRLILCSITGTSIP-----AQMEDITGTPTPLTFIDCVSFTTNVSARFMLADCHOL 828
QY 375 KE-ELRSQVIREACITTAIVYK-----TLNKKIDAFQMS--ILEHLINLIONSAYV 422
DB 829 ETVGATQULRE-LICVPYNAKFVVFPAKMPDVASSLRFCMTDVKDKVKTLEQGENFEV 887
QY 423 IASASTIALK-----YIIRYTHAPKLLK-----ITYDTLNOQ-----SKSDIRSTLCEL 466
DB 888 ARSDIEVLBEKPIYVOCYGNLAPLTGGOQLVNFVYSFKENRILPFSIKIRIDTSOEPGR 947
QY 467 MVLLEPEKQTALEARNATV-LRDTL-----KSIDADCDARHNR--YAWAFRRHPE 518
DB 948 LSFLEKPTTGTLGQIVAVCNLITLPAHKETESDQDEIEKTDRROSFAALAKRYSY 1007
QY 519 LADQIVGTLDIAQAALEREHREGGGGCTGTGTAPETRRVSRIGRTPTLOKPTPSM 578
DB 1008 LTF-----PQMIERSTCAT 1021
QY 579 RSISAVDTAAQRAKVAQYTLVSRORRKLCPNNNSNQ-----SMGAAASGSLPRRL 632
DB 1022 RSLP-----TYS--YKPFSTRPYGQMTAPITVPEPAKSGF--TSL 1060
QY 633 NSNGGTPATTP-----GSVTPRP-----RGAGVYSQSGPGRSPSTKLDQYGI--- 680
DB 1061 SSSSNTPSASPLKSINSVSTSPSPKSTLGTSTTSVVSISDVASPIRSFTMSSPIKTV 1120
QY 661 -----GNVYR--GATGAI PKKAGCIPSTASRETSPTSGGGLMKSMYSTG 726
DB 1121 VSQEPYNIQVSSGLARAPVATEATPLK--GLASNSTSKTSPTVTGSLERSST-- 1176
QY 727 AGSRTPERNRPVPSAPARLAQSRHAHTLVGDGDQDPYVSGDYNRSGMGRKYL 786
DB 1177 -----MTF--PAPKSNIMNYSSSLPK-----SITTSAPLISPLK 1212
QY 787 G-----RDESDDISEASVCSERSFDSSTYRGNKSNYSLSGSHRILMSTQRAFFDIE 841
DB 1213 SVSPVKSADVVISAKITMASSLSPPVKOMPQHAVALVNGSISPLKYPSS-----S 1265
QY 842 TIIOCASTHWSERDGLISLTQYIADGKELTQOOLKCVLDMFRMFMPTHTKVTLSFLD 901
DB 1266 TLINGCKAT-----ATIQEKISAKTNSVSVSAATVTEKVFETTTAMPSPPLRS 1316
QY 902 TVTELLIVHANETSRNGSSCLTRLENKLTGDLNMSHMTKWTLOVVEHYFPTLOLKE 961
DB 1317 YVS--AAPSAGQSLRTPASA--LYTSLGSS--ISATTSSVTSSIIIVPVSVVAVLPEP 1370
QY 962 LFRISDSOTPTTKTRIAIIRFLTDLA-----NTYCKSSDPSPSQSACERTVUKLQ 1016
DB 1371 ALKKLFPDENS--FTKSAALSLPIKTLTTEHPQPHFERTSSPVKSLFLAPSAKLSTP 1428
QY 1017 AADQSMELRQANS-----LVALVNLNTPOMTULLADIP-----KVY 1055
DB 1429 SSLSSQELIKDVAEMKEDLMKMTNIILOTVPEEKFPPELPKBERIDDEBFKIVEKVK 1488
QY 1056 QDSAR-----SCISHMRROSQSCNSGANSPPSSPLSSGPKPLQSPVGPASLQS 1107
DB 1489 EDLVVSEILKKDVCVDNKGSPKSPKDKG--HSPEDWIEFSSSEIRRA----- 1536
QY 1108 HHHQLSISTSPRQRSSVEQELFSSGLDQHNIOKTSSEIRHFGQGYOTRALPNGFN 1167
DB 1537 --RQQAASQSP-SLPERVOYAKAKASBKD--YNLTXYIDVLTNDIGSSSLT----- 1583
QY 1168 GHLOY-HDQGOODSCASLSNSKTSQSANITQSNTPESATWRDLRE----- 1215
DB 1584 -NLKTYFEADAKKDEERKRVLPKPAIQAQEKLNMP-PASWRTSTSEKELCMQADSFTCT 1641
QY 1216 -----RTTQAKSPDADAKVITVS---INNAENGELLASINMESEVVALTLT 1262
DB 1642 DTILESPPDFSGHDDKSPPLSDSGFETSEKTPSPAPOSAETTPGPKPLFHEVPIPVITET 1701

```

```

QY 1263 KQPVLLQTSLTNLGICIKGNCELJN-----KHFRS 1295
D 1702 RREVVHVRISYDPSAG-----DVPQTPREBPVSPKSPFMELEPKPTTSSJKEKX 1753
QY 1296 IRRMLNLTAEHTDVVAGLHVLSKINSNKRNNHMFLELILIKIIQCYQHSKEALR 1355
D 1754 AFQMKASSEEDHN-----RVLSKGMKVEKEETH-----TTTTTMMVYHSPGGE 1797
QY 1356 DIDSMIPRAPSPLDLISININPVATNGEPPTNICALIILEVTEHSGEITDHLDIY 1415
D 1798 GASETIEF-----TMSVHDIIMKAFQSGRDS-----KELAGLEHNSA-----V 1836
QY 1416 FPNLARSADDTOSYMKAAVFCIVKLYFLGEEKYKPKLVNPSKRVLLNLYIEK 1471
D 1837 SPDVHKSAAETSAQNAEK-----DNQMKPKLE-----RIEVIHIEK 1872

RESULT 13
US-09-698-295-10
; Sequence 10, Application US/09698295
; Patent No. 6689584
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATORY FACTOR
; FILE REFERENCE: 06501-068001
; CURRENT APPLICATION NUMBER: US/09/698,295
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/J999/02340
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: JAPAN 10/137631
; PRIOR FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 2781
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-698-295-10

Query Match 2.3%; Score 173.5; DB 2; Length 2781;
Best Local Similarity 18.2%; Pred. No. 0.00031;
Matches 275; Conservative 189; Mismatches 487; Indels 559; Gaps 68;

QY 5 KPSPDIDGFIQWPKADMVKVQLADPLVTFLSDPTNSIVCTDMGFLIDGLMPLTGSHPK 64
D 1053 KVSDLASLGQETKTKTKNDFFIDSKLASADDTGLTLCRK-----KLIOESBDT 1105
QY 65 IAKSLAEFSELIKSLAGSDFNAYTATVLPVID-----RLG-----DSRDYREKAQL-LRDL 117
D 1106 IVSSSKSLHSHVSPSTNDRD-----ATPLSRAMDFSGKGCDESENSTLNSDVTISIDS 1162
QY 118 MEHRLVLPQALIDKLATSCFHKHAKAKVEEFLQITVNLHLEYGTQOLSVRYVYIPVICAL 177
D 1163 SEEDMI-----VQNSNESISEOF-----RTREODVEVLEPKCELY 1198
QY 178 -GDPVTNVRREALIOTLVEIYKHVGRLL-----RP-DRRMDVVPASGLAMLEQKF 225
D 1199 SGEISGNCB-----DRLPYKGTENGAKRPSOOKKLEBRPVAKCS----- 1237
QY 226 DQVKEGELLPSALNNT-----NGNGV-----GLDEADNIGLEREPTRMK 266
D 1238 DQIK-----LAKTTDKNNENRESEKQGTSTFQINGKDNPKIYLGCECLKEIS 1288
QY 267 RPLHSAVSSSLRPK-PNNDVTGD-----AGATMESFESFEVVPQL 308
D 1289 E--SRVVGAVPEPKVNNINKIIPENDIKSLTVKESAIRFINGVDVIMEDFERNSSSTKS 1346
QY 309 NIFAKMDMDIKQVLLVII-----SDKN-----ADMEKKVADLKKIR 345
D 1347 HLLSSDAEGNVRDSELTLPSTKESDSTOTTPPASCPESNSVNOVEDEMEIFTSVKV- 1405
QY 346 ALLLISYHTPOQFVAVOLKELSLSPVD-----ILKEELRSQVIREACT--TIA 391

```

```

D 1406 -----TSPPITSEESNLSNDFIDENGLPINKNNVNGESKRKVTITEVTWTSTVA 1457
QY 392 YMSKTL--RNLDAFCWSLIEHLINLIONSAAVIASTIALKXY-----IK 436
D 1458 TESTKIVKEKGDK-----QTVVSTENCKASTVTTTTTIVTKLSTSTSGSVDIISVK 1511
QY 437 YTHAPKLKIKYTDLT-----NOSKXDIRSTCELMVLF-----E 472
D 1512 EQSKTVATVTTVTDGLTTGGTLVSMYVSKY-STRDKVKLMKRSRPKTRSGTALPSYR 1570
QY 473 EMQKALERNATV-LRDLKKSIGDADCDARHRY-----AYAF 512
D 1571 KFWKSTKKSIFVLPNDLKKARKGGIREVYPYNNAKPALDIWPYSPPTFGITWRY 1630
QY 513 RRRHPE-----LADQYGLT--DIAQALREREGGGGGTGTGTAPET----- 557
D 1631 RLQTVKSLAGVSLMLRLIMASLRMDWAA-----KVPBGGSRTRETSETTEITTEI 1683
QY 558 -----RRVSRIGRTPGTLOKPTPSMRKISA----- 583
D 1684 KRDPVPGYGRFECYCIKTIICPIGVPEPKETPTPORKGLSSALRPKRPETPKQTCGVI 1743
QY 584 VDT-----AAQRAKVAQYTLYSRQKRPPLGPNNSNOM 618
D 1744 IETVVAEBELEWEIRAPAEVEKEKAQAVEQAKKRL-----QOKPVIATST-TSP 1796
QY 619 TGAASGSLPRRLNSNGGTPATTPGSVYPRPRGRAGVSGOSGSRSTSTKLARDQYG 678
D 1797 TSTSTTSTSPAKV-----MVAPISQSVT--TGTMMVLTTKVGSPIATVTFQOKNKHQ 1847
QY 679 GIGNYVR-GATGA-----IPKASG-IPRSTASRET-----SETRSG-GGLM 718
D 1848 TFAFMVAKQGSNSGVNOQKVLGITIPSTGTSOOTFTSPQRTATVTRINTSGSGGTT 1907
QY 719 KRSMTSTGASR-----RTPERNNP-----VRSPARPLAQSBAEHTLGVG 761
D 1908 SNSQVITGPQIRPMTVIRTPLOQSTLGKAIIRTPVWQPAPOQVMTQIIR----- 1959
QY 762 DDGQPDVYSGDYMSGGKRMGRKLMGRDESDPIDSEASVCSERSPDSYTRGNKSNYSL 821
D 1960 -GQP-----VSTAVSANITYSS-TPQKSLTSA 1985
QY 822 SGRHTRLDWSTORAPFDIETIIQFCASTHSEKDGILSLTOYADGKELLTQ-----QQ 876
D 1986 T-STSNLQSSASQPP-----RQQGVKLT--MAQLTQLTQGHGNGQ 2025
QY 877 LKCYLDMFRKKFMDDTHKVTISFLDYTTELIVHANETSRNGSSCLTRLNKLGTDLN 936
D 2026 LTVVIQGGQ-----TTGQLQLIPQGVTVLPGPQ----- 2055
QY 937 SMHSKIMKTLQVNAHEFPYTOQLKELFRISDSQOTPTTKRIALRLFLDLANTYCKSS 996
D 2056 -----OLMOAMENGTVQRFLETPLATTTASTIT----- 2086
QY 997 DFPDSQACACERTYLKLAOLADOKSMELRSQARSCLVALYNLTPQWTLTLADLPKYQ 1056
D 2087 -----TIVSTTAAGTEQROSKL-----SPQW-----QVHQ 2112
QY 1057 DS-----ARSCTHSMRQOSQNSGAN-SPSSSLSSSPKPLQSPSYGPPASLOSH-- 1108
D 2113 DKTLPPAQSSEVGPAAQOPOTAPQSPAPQOTQOSAPQEPVQOTQOTQVSSHPS 2172
QY 1109 -----HHQLGISSTSPSR-QSSVGEOLFPSELDIHNIO-KTSEERHCFGGQVOTAL 1161
D 2173 EAQPTHAOSSKPYVAAQSQPOSNVQOSPVAVGSPQTRIRPSTPSQLSPOQSQOVQTT 2232
QY 1162 A-PNGFNCH--LQYHDQGOQDSCASLSNSKTQSSANT-----TQSNTPESA 1205
D 2233 SQPIPIQPHSTLQIPSGQPOQSPQVQOSTQJLSSGGLNVQSVSPSRPQLQIQOPQ 2292
QY 1206 TMRLDNIRE 1215
; : : : :

```

Db 2293 VIAVPOLOQ 2302

RESULT 14
US-09-698-295-1
Sequence 1, Application US/09698295
Patent No. 6689584
GENERAL INFORMATION:
APPLICANT: Jones, Michael H.
TITLE OF INVENTION: TRANSCRIPTIONAL REGULATORY FACTOR
FILE REFERENCE: 06501-068001
CURRENT FILING DATE: 2000-10-27
PRIOR APPLICATION NUMBER: PCT/JP99/02340
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: JAPAN 10/137631
PRIOR FILING DATE: 1998-04-30
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2907
TYPE: PRT
ORGANISM: Homo sapiens
US-09-698-295-1

Query Match 2.3%; Score 173.5; DB 2; Length 2907;
Best Local Similarity 18.2%; Pred. No. 0.00033;
Matches 275; Conservative 189; Mismatches 487; Indels 559; Gaps 68;

Qy 5 KPSLDGFIQMPKADMKVQVQLAEDLVFLSDDTNSICTDMEFLIDGLMPWLTGSHFK 64
Db 1179 KVSPLASGQEPKSKTKGNDFIDSKLASADIGTLICNKK-----KPLIOESDT 1231
Qy 65 IAGSLKFSLEIKRLGSDFNATVTLPHVID---RLG---DSRDTRKAQLL-LBDL 117
Db 1232 IVSSSKSLHSSVPEKSTNDRD--ATPLSRAMDFGKLGCSSESSTLSSSDTVSIDS 1288
Qy 118 MEHRLVLPQALIDKLATSCFKHKNAKVREELQITVNLHEYGQQLSVRYIIPVCALL 177
Db 1289 SEEDMI-----VQNSNESISEQF-----RTRQGVVLEPLKGLV 1324
Qy 178 -GDPTVNRRAIQTVEIYKHVGRLL-----RP-DLRMDVPAKMLAMEQKF 225
Db 1325 SGESTGNC-----DRLPVKTEANGKKPSQOKLEERPVNKC----- 1363
Qy 226 DOVQOEGLLPSALKNT-----NGNGV-----GLDADNIGLRERTMIX 266
Db 1364 DQIR-----LKNITTDKNNENRESEKKGQRTSTFOINGKDKPKIYLGCEKLEIS 1414
Qy 267 RPLHSAVSSSLRPK-PNVNDVTGD-----AGAVTMESFESSFEVFPQL 308
Db 1415 E-SRVSGNVEPKNNINIKIIPENDISLTVKESAIRPFGINDVIMDFERNSETKS 1472
Qy 309 NIFAKMDMDIYKOVLVII-----SDKN-----ADMEKRVDAKKIR 345
Db 1473 HLLSSDAEGNVRDLETLPTKESDSQTTPSPASCPESNSVNOVDEMEIETSEVKV- 1531
Qy 346 ALLIISTYTOFOFVAVOLKELSLFDV-----ILKEELSOVIREACI---TIA 391
Db 1532 -----TSSEPTSEESNLSNDFIDENGLPIKNENNVNGSKKKTIVTEYTTMTSTIVA 1583
Qy 392 YMSKTL--RNKLDAFCMSILEHLINLIONSAKVIASTIALKYI-----IK 436
Db 1584 TESKTVIVKEQDK-----QIVVSTENCKASTYTTTTTYTKLSTSTGSSVDIISVK 1637
Qy 437 YTHAPKLKIYDTL-----NQSCKDIRSTLCMLVLLF-----E 472
Db 1638 ESKSTVVTVTVDLSLTGGTLVTMTVSKY-STRDKVLMKFSRPKTSRGALPSYR 1696
Qy 473 EMOTKALERNATVL-RDLTKSIGADADCARHSKY-----AYNAF 512
Db 1697 KFTVSTKTSIFVLBNDDKLKLARKGIREVPYFNNAKPAIDWPYSPRPTFGITWRY 1756

Qy 513 RHPPE-----LADQIYGLT---DIAQRALEREGGGGGGTGTGTGAPET----- 557
Db 1757 RLQTVKSLAGVSLMLRLWLSLRMDWA-----KVPPGGSTRTETSETETITTEII 1809
Qy 558 -----RRTVSRIGRTPGTLOKPTSMRSISA----- 583
Db 1810 KRQDVGYGIRFEXCIKRIICPIGVPEPTPEKTPPQKGRSSALRPKRPEPKQTGPVI 1869
Qy 584 VDT-----AAQRKRAVQYTLVSRQKRLPGLPNNNOASM 618
Db 1870 IETVABEELWEIRAFARVEKEKAQAVEQAKKRLF-----QCKPVIATST-TSP 1922
Qy 619 TGAAGSGLRPRLNSGSGTPATPPSYTPRRGRAGVGSQSGSSTSESTLRLRQYG 678
Db 1923 TSSSTSTISPAOKV-----MVAPISGSVT--TGTRKVLTKVGSBATYTFQOKNPHQ 1973
Qy 679 GIGNYR-GATGA---IPKASG-IPRSYASSRET-----SPTRSR-GGLM 718
Db 1974 TPATWVKGQSGNSGVVOYQKVLGIIPSSTGTSQQTTSFQPRATVYIRNTSGSGGT 2033
Qy 719 KRSMTSTGASR-----RTPERNNP-----VRSPAPRLIAQSHAEHTLGVG 761
Db 2034 SNGQVITGPQIRPGMTVIRTPLOOSTLGKAIIRTPVWVGAPQOAVTQILR----- 2085
Qy 762 DDGQPDYVSGDYNRSGGMRKMLMGDESDDIDSEASSVCSERSFPOSTYRGNKSNYSL 821
Db 2086 -GGP-----VSTVVASAPNTVSS--TPGQKSLTSA 2111
Qy 822 SGSHRLDWSTQAPPDIFITIIQFCASHTWSEKQDLSITQYLAQKELYQ-----QQ 876
Db 2112 T-STSNIOSSASQPP-----RPOQGVKLT--MAQLTQLTQHGNGOG 2151
Qy 877 LKCVLDMFRKMFMDTHTKVYSLFLDVTLEILVHANETSRNGSSCLTRLFNKLGTDLLN 936
Db 2152 LTVIIGQGGQ-----TTGQLDLIPQGVTLVPGPQ----- 2181
Qy 937 SMASKIKTTIQVHAETFPQOLKELFRIIISDSQTPTTKRIALRFLDMLANTYCKSS 996
Db 2182 -----OLMOAMPNGTVOGRFLPTPLATTAATTSITT----- 2212
Qy 997 DFPESDQACERTYLKLAQLAADOKSMELRSQARSCIALVALYNLTPOMTLLADLPKYQ 1056
Db 2213 -----TVSTTAAGTEGROSKL-----SPQK-----QVHQ 2238
Qy 1057 DS--ARSCISHMRQSGSCNSGAN-SPSSSEPLSSSPKLPSPVSPASLOSH-- 1108
Db 2239 DKTLPPAGSSVGPAXKQPPQAPGASARPQOTQPSAPQEPVQTOPEVOTQTVSSHVS 2298
Qy 1109 -----HHQLSISTSPSR-OSSVQELFLFSSELDIQHNIQ-KTSEELRHCFGGQYQYAL 1161
Db 2299 EAQPTHAQSSKPCVAAOSQPSQVSNVQSGPVRAVQSPQTRIRPSTPSQLSPGQSQOVQTT 2358
Qy 1162 A-PNGFRGH--IQVHDGQGDSCASLSNKSQTSSANT-----TOSNTPESA 1205
Db 2359 SQPIPIQPHTSLSLQPSGQPSQPSQPOVOSTOTLSSGOTLNAVSVSSPSRPLQIQPQPO 2418
Qy 1206 TMRLDNLERE 1215
Db 2419 VIAVPOLOQ 2428

RESULT 15
US-09-949-016-11433
Sequence 11433, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMERHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: US/09/949,016
PRIOR APPLICATION NUMBER: 60/241,755


```

PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11433
LENGTH: 2733
TYPE: PRT
ORGANISM: Human
US-09-949-016-11433

Query Match      2.3%; Score 172; DB 2; Length 2733;
Best Local Similarity 17.3%; Pred. No. 0.0004;
Matches 270; Conservative 245; Mismatches 562; Indels 488; Gaps 64;

QY 71 EAFSELIRKLGSDFNATYATVLPVHIDRLGDSRDYREKAQLLDLMEHRLPPOALID 130
DB 562 EGFQALVKQM-----NQLDQKTNOI--DLQAEISENQAIHQ 597
QY 131 KLAIS-----CFKHKAKVREEFLOITIVALLHEGYTOOLS 165
DB 598 KLITNTDASDSDSVALVETIVISPTCTGSEHMKPELEKIL-----AL-EKEKEDQ 651
QY 166 VRVYIPVYCALLGDPVTNVREAIIQTLVEIYKHVGDRLPDLRMDVYPAKSLAMLEOKF 225
DB 652 KKL-----QELATSRKAILKKAQEKERH-----LREELKQCKD-----DYNRLQEQF 693
QY 226 DOVKQEGILLPSALK-----NTNGNGVGDDEADNT-----GLAEPRPMKIR----- 267
DB 694 DEOSKENENIGQLPOLQIÖVESIDGKLPTDQOESCSTGPLEPFEKATEQHHTQPV 753
QY 268 -----PLHSVSSSLRPKPNVDVYG-----DAGAVTME-----SFESS 301
DB 754 LESNLCPRMPSHSEDASALOGSTVAQIKAKLEIHAKEVLEELKAVSTTSLETUKSEV 813
QY 302 FEVVPOLN-----IFHAKMDDIYKÖVLV 325
DB 814 FOLQOINKQGLIEISLKTVSHEAEVHAESLQOKLESSOLQIAGHEHLEIOPKLEDEK 873
QY 326 IISDKRAD-----WEKVDALKIKRAL-----LILSYHTQPOFVAVQ-----LKEJSL 368
DB 874 LISKEEDVSYLSGOLSEKEAALITKIÖTEIIEQEDLIKALHTQLEMQAENHERIKÖQV 933
QY 369 SPVDI-----LKEELRSQVIREACTITAYMS-----KTLRKMLDAFCMSIEHIN 414
DB 934 ELCEMKÖKPEEIGESRAKÖQIÖRKIÖALLISRKALKENKSLQÖEL--SLANGTIERLTK 992
QY 415 LI-----ONSAKVIASTIALKYIIKYTHAPKLIKITYDTLNSKSDIRSTLCE 465
DB 993 SLADVESQVSAQNKER-----DTVIGRLALLQÖERDKLITEMDRSLLENQÖS---LSSSE 1044
QY 466 LMTLLFE-----EWÖT--KALERNATVLRDTLKKSIGAD 498
DB 1045 SLKTLLEGITEDKEKLVKEIESLKSKIAESTWÖKHEKLEKÖEYIILQÖSEYENSVNEAF 1104
QY 499 C-----DARRHSRYAYMAFR-----RHPPELADÖIYGLDIAÖRAL 535
DB 1105 RIQHVVEAVRÖKÖELIYKGLRSTEAANKETEKQLOEADÖEMÖKÖKPKSKÖQÖKIL 1164
QY 536 EREREGG-----GGGGTG-----TGCTABETRTVSRIGRTPGTLOKPTPSMRISIA 583
DB 1165 ELEENDRLRAEVHPAGDTAKCEMETLSSNSAMKEELERVVMEYETILSKKFQSLMSEKD 1224
QY 584 VDTAAQAKVAQÖTTLVSRÖRKPLGPNNSNQASMTGAASGLPRPRLNSNGSGTPATT 643
DB 1225 SLSEEVÖDKHÖIEGVNSKÖANLEATERKHÖNÖTNYT--EEGTÖSIP-----GETEÖD 1275
QY 644 PÖSVTPRPRGRAGVÖSÖGPRSTÖPS--TKLRDÖYGGIGNYRGATGAIÖPKASGIPRST 702
DB 1276 SLISMSTRPI-----CSESVPASAKSANPAVSKDÖSSHDEINNYLQ--QIDÖLKENIAGLEBEK 1330

```

```

QY 703 ASRETPSTRSGGIMKRSMTSTGAGSRRTPERNNPVPAPAPALLAÖREAEHTLGVCD 762
DB 1331 ÖKNKEFSQTLLENKNTLLSÖISTÖDÖELKMLQÖE-----VTRKMLNÖQIÖEELS----- 1380
QY 763 DGÖPDYVGDVYMRSGMRGR--KLMGRDESDIDSEASSVCSERSFDSÖYTRGNKSNYSL 821
DB 1381 -----RVYTKETABEKKDLEERLMLNÖLAE-----L 1407
QY 822 SGSHTRLDWSTÖRAPFDIETIIOFCASTHWSERKDGILSLÖYIADGKELTÖOQKCV- 880
DB 1408 NGS-----IGNYÖDÖVDTAQIKNELLE-----SEKKNLK-----KCVS 1440
QY 881 -LDMFRKMFMDTRKVTSLFIDVTTELILVHANTSRNGSSCUTRLE-----NKLGT 932
DB 1441 ELEBEKÖÖLVÖKTEKVSERKEVLEKIOGAKÖKPGNKSNAKELQÖELLKÖQÖEVKÖQÖK 1500
QY 933 DLNLSMH--SKIWTTLOVHEYPTÖLQÖKEL-----FRITS 967
DB 1501 DCIRYÖEKISALEETVAKL--EFVÖTESQ--KÖLETTKENLAÖAVEHRRKKAQÖELASFYKVL 1558
QY 968 DSTÖPTTKTRIALIRFLDTLANTYCKSDPDSQÖACERTVYKLAQÖADQÖKMEIERS 1027
DB 1559 DDTQSEARVYLDNLKIKKELQÖS--NKESVYÖMQKÖEDLERLEQ--AEKÖLKEKK 1613
QY 1028 QASÖCLVAL-----YNLNTPÖMTLLADLPKYVÖDSASCSHSHRQÖSÖGNS--- 1076
DB 1614 NMÖEKTLALREKVLHEETIGELIÖVTLNKXÖ-----KEVÖQÖLÖENLDSVT 1659
QY 1077 --GANSPPSSPLSSSPKLPQÖPS--VGPFY--SLÖSHHÖLÖSI-----SSTPSRQÖSV 1126
DB 1660 QLAÖFTYMSLÖDÖDRVIDEAKWÖKRSÖDALSÖEERIRLÖEDNCSVLKÖDÖRQMSI 1719
QY 1127 EÖELLFSSÖLDIÖHNIO-----KTSEIRHCFGGÖYÖTALAPNGFNGHLQYHDÖG 1176
DB 1720 HMEELKINISRLÖHKQÖWESKQÖTEVQÖKQVCDTQÖGENKELLSQÖLETRHL--YH--S 1776
QY 1177 QÖDSÖASLSSNSKQÖSSANTQÖSTPESATMRNLNLE-----RERTONASPTDDAV- 1230
DB 1777 SÖNLAKLESÖLKSÖKDÖLTDLSNLSÖKÖKNGNLGÖIIRQÖADIQNSKFYÖQÖLTD 1836
QY 1231 -----ITVSINNAENGELILASNLMESEVYVALTLTKQÖPVELLQÖSLTNLÖIC 1280
DB 1837 LQASRELTSLRHEINKEÖKÖTISLÖSG--KEKALQÖVAL----- 1873
QY 1281 IKGNCÖELPNKGFRSINRMLNILEAHTDVYIAGHLVLSKIMSNOGRHNMHFLÖLIL 1340
DB 1874 -----AELRQÖHÖKEI--KELENLLSÖEBEENIVLEENKKAVDKTNQÖL---METLKÖTIK 1923
QY 1341 LKTIQ 1345
DB 1924 KENIQ 1928

```

Search completed: November 23, 2005, 15:13:24
 Job time : 70 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 23, 2005, 15:06:50 / Search time 10 Seconds

(without alignments)
452.221 Million cell updates/sec

Title: US-10-030-850-2

Perfect score: 7611

Sequence: 1 MAYRRPSDLGFTIQMPKAD.....RNCISGGSGSTKNSSAASSS 1492

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 17545 seqs, 3030971 residues

Total number of hits satisfying chosen parameters: 17545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications AA.New:*
1: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	191	2.5	4384	1 US-10-821-234-1120	Sequence 1120, Ap
2	186.5	2.5	761	1 US-10-485-517-252	Sequence 252, App
3	163.5	2.1	1448	1 US-10-485-517-212	Sequence 212, App
4	153.5	2.0	1402	1 US-10-971-982-2	Sequence 2, Appl
5	146.5	1.9	3144	7 US-11-055-035-1	Sequence 1, Appl
6	146	1.9	5024	1 US-10-793-626-2964	Sequence 2964, Ap
7	133	1.7	594	1 US-10-131-826A-10	Sequence 10, Appl
8	131.5	1.7	482	1 US-10-821-234-972	Sequence 972, App
9	125	1.6	1463	1 US-10-971-982-3	Sequence 3, Appl
10	124	1.6	1618	1 US-10-984-645-2	Sequence 2, Appl
11	123	1.6	1451	7 US-11-046-346-1	Sequence 1, Appl
12	122	1.6	1189	7 US-11-074-176-134	Sequence 134, App
13	121.5	1.6	1290	1 US-10-485-517-141	Sequence 141, App
14	119	1.6	2897	1 US-10-499-715-2	Sequence 2, Appl
15	116.5	1.5	1377	1 US-10-821-234-1070	Sequence 1070, Ap
16	114	1.5	1493	7 US-11-004-057-4	Sequence 4, Appl
17	110	1.4	1279	1 US-10-793-626-3188	Sequence 3188, Ap
18	108.5	1.4	636	1 US-10-485-517-170	Sequence 170, App
19	108.5	1.4	1637	1 US-10-821-234-1204	Sequence 1204, Ap
20	108	1.4	1125	1 US-10-821-234-1444	Sequence 1444, Ap
21	108	1.4	1493	7 US-11-004-057-21	Sequence 21, Appl
22	108	1.4	2053	7 US-11-013-759-9	Sequence 9, Appl
23	106.5	1.4	313	1 US-10-512-184-72	Sequence 72, Appl
24	106.5	1.4	576	1 US-10-512-184-65	Sequence 65, Appl
25	106.5	1.4	625	1 US-10-512-184-47	Sequence 47, Appl

26	106.5	1.4	1207	1 US-10-821-234-1109	Sequence 1109, Ap
27	105.5	1.4	748	1 US-10-821-234-888	Sequence 888, App
28	105.5	1.4	989	1 US-10-821-234-975	Sequence 975, App
29	105	1.4	258	1 US-10-512-184-26	Sequence 26, Appl
30	105	1.4	327	1 US-10-512-184-62	Sequence 62, Appl
31	105	1.4	327	1 US-10-512-184-64	Sequence 64, Appl
32	105	1.4	327	1 US-10-512-184-63	Sequence 63, Appl
33	105	1.4	589	1 US-10-821-234-1494	Sequence 1494, Ap
34	105	1.4	868	1 US-10-821-234-1082	Sequence 1082, Ap
35	104.5	1.4	251	7 US-11-054-515-1810	Sequence 1810, Ap
36	103.5	1.4	509	1 US-10-821-234-1093	Sequence 1093, App
37	103.5	1.4	514	1 US-10-821-234-998	Sequence 998, App
38	103.5	1.4	604	1 US-10-793-626-390	Sequence 390, App
39	103	1.4	261	1 US-10-512-184-35	Sequence 35, Appl
40	103	1.4	643	1 US-10-510-386-8	Sequence 8, Appl
41	102.5	1.3	252	1 US-10-512-184-28	Sequence 28, Appl
42	102.5	1.3	1107	1 US-10-485-517-145	Sequence 145, App
43	102	1.3	1970	1 US-10-821-234-1641	Sequence 1641, Ap
44	101.5	1.3	1187	1 US-10-821-234-955	Sequence 955, App
45	101.5	1.3	1560	7 US-11-059-982-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1	US-10-821-234-1120	Sequence 1120, Application US/10821234
Publication No.	US20050255114A1	
GENERAL INFORMATION:		
APPLICANT:	Labat, Ivan	
APPLICANT:	Stache-Crain, Birgit	
APPLICANT:	Andarmah, Susan	
APPLICANT:	Tang, Y. Tom	
TITLE OF INVENTION:	Methods for Diagnosis and Treatment of Preeclampsia	
FILE REFERENCE:	821A	
CURRENT APPLICATION NUMBER:	US/10/821,234	
CURRENT FILING DATE:	2004-04-07	
PRIOR APPLICATION NUMBER:	US 60/462,047	
PRIOR FILING DATE:	2003-04-07	
NUMBER OF SEQ ID NOS:	1704	
SOFTWARE:	pf_seq_genes Version 1.0	
SEQ ID NO	1120	
LENGTH:	4384	
TYPE:	PRT	
ORGANISM:	Homo sapiens	
US-10-821-234-1120		
Query Match	2.5%; Score 191; DB 1; Length 4384;	
Best Local Similarity	18.1%; Pred. No. 0.00016;	
Matches	303; Conservative 242; Mismatches 613; Indels 518; Gaps 76;	
QY	69 SLEAPSELKRLGSDFNATATVLPVIVDRIGDSRDTREKAKQLLRDLMEHRVLPQAL 128	
DB	913 SLGRASASLRPSSD-RSYT-----LNSSYARDSMIE-ELVPSKEQLLPTTRFD 963	
QY	129 IDKL-----ATSCFGRKN--AKVREEFLQI-IVNAL--HEVGTQQLSVRYIYP----- 172	
DB	964 SDSLRYHWMADTLDNVLVSPHSGFLVAFMVDARGSMGSRHHGRILIIIPRCKTA 1023	
QY	173 ---VCALL-----GDTVANREALITQVIEIYHGVDRLLRPDLRRMDVP----- 214	
DB	1024 PTRITCRIVKHKLANPPHERRGISRLVEMGPAGQFGLPVI---VEIHFSGMRCK 1080	
QY	215 ASKLAML-----EOKFDVQKOEGLLPSALKNTNGNGVGLDEADNIGIRERTMI 265	
DB	1081 ERELVLVSENGETKKEQFDS-KREDL-----TELLNGMBEIDSPBELG-KKRICII 1133	
QY	266 KR--PLHSAVSSSLRPKN-VNDVTGAGAVTWSEFSSP-----EVVQ 307	
DB	1134 TKDFQYFAVVSRIKQESNQIGPEGILSTTVPLVQSGFEGALTKRIRVGLAQAPV- 1192	
QY	308 LNFPAKMDIYKQVL-----VIISDKNADMEKRV----- 338	

```

Db      1193 -----DEIVKILGNKATFSPIVTEPRRKRKHKEITMTIIVPPSPGSGVNGYKG 1243
Qy      339 DALKIRALLISYHTPOFOVAVOLKELS-----LSFV-----DIL 374
Db      1244 DTTNRLILCSITGGSP-----AQMEDITGTTPLTIKOCVSTTVNSARFPLADCHOVL 1299
Qy      375 KE-ELRSQVIRACITIAVMSK-----TLRNKLDAPCWS--ILEHLINILIONSAYV 422
Db      1300 ETVGATQGLYRE-LICVPMKAFVFAKNNDPVESLRCFCMTDDKVDKDTLQOENFEVB 1358
Qy      423 IASASTIAK-----YIKYTHAPKLIK-----YTDTLNQ--SKSKIRSTLC 466
Db      1359 AASKDIEVLEGKPIYDCYNLAPLTKGQOLVFNYSFKENRLPFSIKIRPTSGPCPR 1418
Qy      467 MWLFEEMOTKALERNATV-LRDTL-----KKSIGDCCDARRHSR---YAWAPFRHPE 518
Db      1419 LSFLEPKTKGLPOTAVCNLNTLPAKKETESDODDEIEKTRROSFASIALKRYSY 1478
Qy      519 LADQIYGLDIAAORALEREREGGGGTGTGTAPETRTVSRIGRTPTLOKPTPSM 578
Db      1479 LFE-----PGMIERSTGAT 1492
Qy      579 RSISAVDTAAQAPAKVRAOYLYLSRORPLGPNNSNOA-----SMTGAASGSLPRRL 632
Db      1493 RSLP-----TYYS--YKPFSTRPYOSWTAPITVPGAKSGF--TSL 1531
Qy      633 NSNSGTPATTP-----GSVTPRP-----RGRAGVSOQSGSTSPSTLRAQYGI--- 680
Db      1532 SSSSNTF8ASPLKIMSVSTSPFKSTLGASTTSVKSISDVAPFIRSLRMTSPKITV 1591
Qy      681 -----GNYYR--GATGALPKKASGIPRSTASSRETSPTRSGGGLMKRSMYSTG 726
Db      1592 VQSPYNIQVSSGTLARAPAVTEARPLK--GLANSTFSRSTSPVTTAGSLERSST-- 1647
Qy      727 ACSRTTPERNVVRSPARLLAOSREAHNTLGVDDGDPYVSGDYMRSGMGRKXLM 786
Db      1648 ---MTP---PASPKSNINMSSSLPFK-----SITSAAPLLSSPLK 1683
Qy      787 G-----RDESDIDEASVCSERFDSYTRGNKNSYLSGSHRLMDSTORAPFDLE 841
Db      1684 SVSVYKRVAVIYSAKITMASLSLSPVKOMPGHAEVALVNGSIFPLKASS-----S 1736
Qy      842 TIIQCASTHMSERKGLISLTQYADGKELTQOOLKCVLDMFRKFMDTHTKVSLEFD 901
Db      1737 TLINCKCAT-----ATLQEKISSATNSVSVSAATDVEKVFSTTAMPSPSIRS 1787
Qy      902 TVTELILVHANETSRRNGSSCLTRLENKLGTDLNSHMSKIWTQVYHEVFTLOLKE 961
Db      1788 YVS--AAPSAPQSLRTPSASA---LYTSLGSS--ISATTSSVTSSITIVPVSVVAVLPEP 1841
Qy      962 LFRISDSTQPTTTRTIALILFLDLA-----NRYCASSDPRPSQOACERTVUKLOL 1016
Db      1842 ALKKLPDUNS--FTYSAAALSPITKLTTHTPQHPFSRTSPVKSFLAPLSALKLTP 1899
Qy      1017 AADQSMELRSGQARSC-----LVALLYNLTPQMTLLADLP-----XYV 1055
Db      1900 SSLSSSQGLIKVAMKEDLMMTAILLOTDVEKPEPQELPKERIDEBEFKIVENK 1959
Qy      1056 ODSAR-----SCISHMRQOSQNSGANSPPSSPLSSSPKPLSPVSGVPASLOS 1107
Db      1960 EDLVVSEILKKDVCVDNKSGSPKSPDKG--HSPEDDMIEFSESSEIREA----- 2007
Qy      1108 HHHOLISSTSPRSQSSVEOELLPSSELDIOHNIOKTSEIRHCFGGQYOTALANGFN 1167
Db      2008 --RQQAASQSP-SLPERVQVAKAASEKD--YNLTKYIDVTNIGSSSLT----- 2054
Qy      1168 GHLQY-HDQGOODSCASLSSNSKTOS SANTTOSNTPEGATMRLDLRE----- 1215
Db      2055 -NLKFKFEDAKKDGQKRVLKPAIALQEHKLKP--PAMWRTSTSEKELCMADSFPGT 2112
Qy      1216 -----RTTQAKSPTDAKIVYS--INMAENGELILASNLMESEVVRALTLT 1262

```

```

Db      2113 DTLESPPDSQHDQKSPSLSDSGFETRSKTPSAPQSAETTPKPLFHEVPIPVITET 2172
Qy      1263 KDQVVELLQTSILTNIGLICIKGNCGLPN-----KHPRS 1295
Db      2173 RTEVHVHVRISYDPAG-----DVPOQDEBPVSPKSPPTFMLEBKPTTSSIKENK 2224
Qy      1296 IMRMLNLLEAHTDVVIALHVLSTKIMRSNKRNNHMFLELLIKTIQCYOHSKALR 1355
Db      2225 AFQKASSEEDDH-----RVLSKGMRVKEETH-----ITTTBRVYHSPQGE 2268
Qy      1356 DIDSMPILASLPUDLSINIVNPIVANGFEPTNLCAKILLEVTEHHGSEITDAHLIV 1415
Db      2269 GASERIEE-----YMSVDHIMKAPSGRDP-----KELAGLFHKSA-----V 2307
Qy      1416 FPNLARSADDTQSMWRKAACVIVLYFVLGEEKVPTLASYLNSKVALLVYIEK 1471
Db      2308 SPDVHKSALETSAQHAER-----DNQMKPLLE-----RIIEVHIEK 2343

RESULT 2
US-10-485-517-252
; Sequence 252, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynex Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic polypeptides
; FILE REFERENCE: P1006290
; CURRENT APPLICATION NUMBER: US/10/485,517
; PRIOR FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 252
; LENGTH: 761
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-485-517-252

Query Match      2.5%; Score 186.5; DB 1; Length 761;
Best Local Similarity 18.3%; Pred. No. 2.5e-05;
Matches 136; Conservative 129; Mismatches 358; Indels 119; Gaps 19;

Qy      545 GGTGTGTGTAETRTVSRIGRTPTLOKPTPSMRSISAVDTAAQAKVRAQYTLYSRQ 604
Db      81 GSTSTLSNSTRSGSTIS-----TSTISBSTSTKSSSVSTLSLMSSTSTSLSDSTSLSTS 136
Qy      605 RKPLGPNNSNOASNTGAAAGSLPRPLNLS-----GCTATPTGATTPPRGRA----- 655
Db      137 LSDSTSDSKSLSLSTMSSTSDSTSTSKSDSISTSTLSLGSSTSESDSTSSSEKSDSTS 196
Qy      656 ---GVSQOPCSRSTSPSTKLRDQYGINYRGATGAIPKKASGIPRSTASRSTSPTR 712
Db      197 MSIMSOSTSGSTSTSTSTSLSDS-----TSTLSLSASANO 233
Qy      713 SGGGLMRSMYSTGAGRRTPERNNYPVPSAPARLLAOSREAHNTLGVDDGDPYVSGD 772
Db      234 SGVDSNAGSAGANSSTSTSTSESDQSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 290
Qy      773 YMRSGMRMRKXLMGRESDDIDEASVCSERFDSYTRGNKSNYSLSG----- 823
Db      291 TSGGSAVSTASLSGSESDSOSISTASBSTSEASTSLSDSTSTSTSTSTSTSTSTSTSTST 350
Qy      824 -----SHTRLDST---ORAPFDIET--IIQCASTHMSERKGLISLTQYLA 867
Db      351 NSASASDSLSSTSLSDSTASMSQSSSDSOSTSASLSDSLSTSTSNRMSTIASLSTSVS 410
Qy      868 DGKELTQOOLKCVLDMFRKFMDTHTKVSFLDPTVELILVHANETSRRNGSSCLTRLF 927

```

```

Db      411 TSESGSTSESTSESTSTSLSDSQST-----SRSTASGASTSTSTS 454
Qy      928 NKLGLDNLNSMNSKSI-----WKTLOVHEHFFPTOLQKELFRILISSTOPPT---TKTRIA 980
Db      455 DSRSTASTSTSMRTSTSDSQSMSTSTSTSTMSDSTSLSDSVSTSTSTSTASTSGMS 514
Qy      981 ILRFLTDLANTYCKSKSDFPS---DOSQACERTVLKLAQLADOKMELRQARSCVLVLY 1037
Db      515 VSISLSDSTSTSTASSEVWSASISDSQSMSESVNSESVSESNSSDSKSMGCTSVS-- 572
Qy      1038 NLNTPQMTLLADLPKYODASRCSHSMRROSGCNSGANSPPSSPLS-----SSSPK 1092
Db      573 DSGSLSVSTSLRKSSESVSESSSLSC-----SQMSDVSSTSDSSSLSVSTSLRSES 624
Qy      1093 PLOGSVGPFAFLQ---SHHQLSTSTSPRSRQSSVEQELLFSESLDIOHIOKT-SPE 1148
Db      625 VSESLSLSDSKSTSGSTSTSTSGSLSTSTSLSGSESVSESTSLSDSISMSDSTSDSDS 684
Qy      1149 IRHCFGGGYOTALAPNGFNGHLQYHDQGGQDCASLSNSKTOSSANTQSTPESATWR 1208
Db      685 LSGSISLSGSTSLSTSD-----SLSDSKSLSSSQSMSESTSTSVSDSQSS--TSNQ 737
Qy      1209 LDNLERTTQNAKSEPTDAKV 1230
Db      738 FDSMSISASESDSMSTSDSSSI 759

```

```

RESULT 3
US-10-485-517-212
; Sequence 212, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biogenex Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629MO
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 212
; LENGTH: 1448
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-485-517-212

```

```

Query Match      2.1%; Score 163.5; DB 1; Length 1448;
Best Local Similarity 20.1%; Pred. No. 0.0022;
Matches 196; Conservative 143; Mismatches 380; Indels 255; Gaps 46;

Qy      82 SFNNATATVAVLPHVIDRLD---SRPTAREKAQULLRLDMERHVLPP-----QALLDK 131
Db      546 NNINATATT-----GDVTTAKTAVOKVO---QLHANFVKKPKAGKELIDQAAADK 592
Qy      132 LATSCEFKHNAKVR--EFLQITIVNALHEVGT--QQLSVRVYIPVYCALLGPTVAVREA 187
Db      593 -KTQLEQTNAOQELINDAKQEVDTLBNAKTNVDSSINEVY-----DNAVEGKA 643
Qy      188 ---AIGTIVEIYK---HVGDRLRPDLRMDVPA---SKLAMLEKFPQVQEGILLPS 237
Db      644 KINAVYTFSEYKKDALAKIEDAYNAKYNADNSNASTSEISIAEKOKLKLQ-----T 697
Qy      238 ALKATNGNGVLD-----EADNIGLAEPRPRMIRKPLHSAVSSSLAPKPVNDVTGD 289
Db      698 ADQNNV-NQATSKDIDEVQIHNDLNDINDYTTPTGKKEASATTDLYAVADQK--NNISAD 753

```

```

Qy      290 AGAVTMSEFSEFVVPQNLTFHA-----KMDMDIKOYVLIIISDNADMEKRV 339
Db      754 TNA-TQBEKQQAIIQVQD-NVQTALESINNGVDGDDALQ-----GKAIDAIOVD 805
Qy      340 ALKKIRALLIISYHT-----QQFVAVOLKEISLSPVDILKEELRQVIREACITIA 391
Db      806 ATVPKANOALEVKAEDTKESIDOSDOLITAEKTEALAMIQIDQAK-QGITDA-TTIA 863
Qy      392 YMSKTLRNKLDAF-----CWSILHLINLIONSKAVIASASTIAIKYIKTH 439
Db      864 EVERAKAQGLAFNFIQIDSTREKOKAIELETALDOIAGVNVNADATTEEKE---AFTN 920
Qy      440 APK-LKITYDTLANSKDIRSTLCBMLLEBEMQTKLERRA-TVLRDTLKKSI-- 494
Db      921 ALEBILSKATEDISDQTNNAIATVKNALQOLKQRIINPEVKKNALAIHEVVKQIEI 980
Qy      495 ---GDADCARHRHRYAWAFRRHPELADQIYGT-----LDIAQALERERE 540
Db      961 IKMADADSAKEIAR---TDLGRYFDRFADKLDKTQTAEVAELQNVITIPAIELVPOND 1037
Qy      541 -----GGGGGGTGTGTGTAPETRTVSRIGR-----TQGTLOKPTPS 577
Db      1038 PDANDTNGIDNNDATANSNANATPENTG-QPNVSETTAN-GKADASPTTPNNSDAATGE 1095
Qy      578 MRSISAVDTAAQAKAQAQYTLYSRQKRPGLPNNNSQASMTGAASGS--LPRRLNSN 635
Db      1096 TTATSATDGA-----NDKPOANNSSVDASTNSTPTMDNVTSKPEVEST 1139
Qy      636 SGGT---PATPGSVTPRPRAGVSOQPSRSTSPSTKLRDQYIGNYRGATGAIP 692
Db      1140 NNGTIDKVTETDNTIP-----AESTTNNSTTTATN-----ENAPGSTATAAP 1183
Qy      693 KASGIPRSTASRRETPRSGGIMKRSMTGAGSRT-----PERNPVR----- 740
Db      1184 TTASTEAASADSKONASVNDNS--KQNAEVNNSAESOSTVDKVAQPKSENKAKAEKDS 1240
Qy      741 PSAIRLLAQRREAHITGVGDQDPVYSGDYMASGGMRGRKLMGDESDDIDSEASS 800
Db      1241 DSTNOSWESSTTELPLPSADITEPNVNTSKD-----KEESTTQTDIACQ 1285
Qy      801 VCSERSPDSYTRGNKSNYSLSGSHTRLDWSTQRPDIETIIFPCASTHMSERKGLI 860
Db      1286 LKSETNVAS-----NEAKSPSKADTEVSNKPSSTASSEAK--EKMTSTVNSQDDTAT 1337
Qy      861 SLTGYLADKELTQOQLKCVLDMPFKMMDTHTKVYSFLDTVTLLILVHAN-ETSRNCS 919
Db      1338 ADTN-----DTQKSVGSAANNKATQNDGANASPATVSNCS 1372
Qy      920 SSCLTRLNKLGT 933
Db      1373 NSANQMDMLNVTMTD 1386

```

```

RESULT 4
US-10-971-982-2
; Sequence 2, Application US/10971982
; Publication No. US20050244899A1
; GENERAL INFORMATION:
; APPLICANT: Rosenfeld, Michael G.
; APPLICANT: Glaes, Christopher K.
; APPLICANT: Rose, David W.
; APPLICANT: Torchia, Joseph
; TITLE OF INVENTION: A Transcription Factor Coactivator Protein, p/CIP
; FILE REFERENCE: 6627-P11021
; CURRENT APPLICATION NUMBER: US/10/971,982
; CURRENT FILING DATE: 2004-10-21
; PRIOR APPLICATION NUMBER: US/09/445,353
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: PCT/US98/12263
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/049,452
; PRIOR FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 14

```

SOFTWARE: Patentin version 3.1
 SEQ ID NO 2
 LENGTH: 1402
 TYPE: PRT
 ORGANISM: Mus musculus
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (3121)..(3121)
 OTHER INFORMATION: "n" is any nucleotide
 US-10-971-982-2

Query Match 2.0%; Score 153.5; DB 1; Length 1402;
 Best Local Similarity 18.6%; Pred. No. 0.0095;
 Matches 276; Conservative 196; Mismatches 557; Indels 451; Gaps 69;

126 QALDKLATSCFKH-----NAKREELQTVNALHEVGTQSL-----V 166
 6 ESSLDPLAESRKRLPCAPGGLVYSGEKWRREDSKYIELAELISANLSDIDNFV 65
 167 RYIIPVPCALLGDPVNVNREAAIQTLVEIKVGDRLRMDVDPASKLAMEQKFD 226
 66 K--EDKAILKETVRQIRQ-----IKEGKTISD--DDVQKADVSTGCGVI 109
 227 QVKGGLLPSALK-----NTNGVGLDEADNIGLRERPTRMKRLHSAVSSLRK 280
 110 DKDSLGLPLLQALDGFLEFVNVRDGNIVFSENVTOYLOKQEDLVNTSVSTLHPRR-- 167
 281 PNVNDVTGAGAVTESFESEFVVPOLNIFPAKMDDIYKQVLIIISPKNADWEKRYDA 340
 168 -----KQFLNTYQ-NPQLMEFLG-----LWRTDKKAPYLIIYRM 201
 341 LKTRALLILSYHTOPQVAVQKELSLSEVDILKE---ELRSQVIREACITAYWSK 395
 202 LMKTH-----DILEDVNASPETRQRYETMQCFALSOBRA 235
 396 TLRLNLDAFCWSILEHLINLIONSAAKVIASASTIALKYIKTHAPKLKIYTDLNGSK 455
 236 MLEBGEEDLQCC-----MICVARRVTPAPPSPPSFTTRHDLGKVVNIDTNSLRSSM 287
 456 SKDIRSTLCMLVLLFEEMQTKALERMATVLRDLTK--SIGDADCDARRHSRYAVMAFR 513
 288 RPG-----FED-----IIRCIQRFPSLNDGGS-----WSQK 314
 514 RHPEPLADQIYGTLDIAQALEREEREGGGGGTGTGTABETRTVSRIGRTPTG--- 570
 315 RHYQGA--YVHGHAETPVYRF-----SLADGTIVSAQTK--SKLFRNVTNDR 358
 571 -----LQK-----PTPSMSISAVDTAAQARAVRAQYLLYSRQRP----- 607
 359 HGFISTHFLQREONGYRNPPIPODDGI-----RPPAAGCGVS 395
 608 LCPNNSNOASMTGAASGLPRRLNSNGCTPATTPGVTPRPRGAGVSOOSGSRST 667
 396 MSPNONGV--MMGSTTYG--VPDPSTGQMGGA-----RYGASSVASLTGQSLQ 442
 668 SPSTKLRLQYQ--GIGNYRGATGAI PKASGI PRSTASRETSPTRSGGLMKRSMYSTG 726
 443 SPSSVONSYSGLSMSSPHGSPGLDPNQNM-----ISPFRNGSPKASHQFPA 493
 727 AGSRTPERRANNVRSAPARLLAQSREAHYL-----GYGDD----- 763
 494 AGA-----HSPMGSGGNTGSHSFSSSSLSALQAI SEGGVSLTSLSPGPKLDNSPMM 547
 764 --GQPDVYSGDYMRSG--GWRMGRKLMGRDESDIDISEASSVCSERSPDSYTRGNKSNYS 820
 548 NISQPSKYSGGDSKSPGLIYC-----EQNPVE--SSVQSNRSDRPQYKESKES-- 594
 821 LSGSHRLDWSTQRAFPDD--LETTIQF--CAST--HMS-----ERKDLGLISTQY 865
 595 -SGEVS-----ETPRGPLESKGKHKLQLLTCCSDRGSHTLNSPLDPKCKSSVSVYS- 648
 866 LADGELTQOQOKCVLDMRKMFMTHTKVYSLFL--DVTVELLVHANENSRNCS--S 921

649 PSGVSSSTGTVSSTSNVHSGLQEKRIILKLLQNGNSPAEVAKITAEATGKDTSTAS 708
 922 CTRLEFNKLGTDLLNSMHSKIMKTLOVVEYF-----PTQLQKEL----- 962
 709 C-----GEGTTRQEQSLSPK--KKNENMLRYLLDRDPSDVIAKELAQADSGSKLSQC 761
 963 ---FRISDSYQTPPTTKRI-----ALRFLTDLANRYCKSSDPSDQACE 1007
 762 SCSTNPSSGQEKDKRIKTENDEVSGDLNDLAILGDLT-----SSDFYNNFTNGCH 813
 1008 RTVLKLAQIADQKSMELRS--QARSCVALAYN--LNTPQMTLLADLPKYODSA--- 1059
 814 PGAKQ--QMFAPGSPSLGIRSPQVQSVPRPNRAVSLDSP---VSVSGPVPKXVNSAPRG 868
 1060 ---RSCISHMVRROSQSCNSGAN-----SPSSSP----- 1085
 869 LPKQPILAGNFRMMDQENYAGANNPVRNVEVNTSPSPGDWGLANSRASMEPLASSPLG 928
 1086 ---LSSSPKPLQSPSVGPFASLQSHHQLSISSTSPRSROSSVEQELLFSSSLDIQH 1140
 929 RTGADYSATLPRPAMGGSV--PTLPLRSNR-----LPGARPSLQOQOQOQOQOQOQO 979
 1141 NIOKTSERIRHCFGQYQOTALAPNGFNGLQYHDQOQDSCASLSNSKXTOSANT----- 1196
 980 QO 1038
 1197 -----TQNTBESATMR--LDNLER--ERTONAKSPDDAKVITVSINMAENGELI 1244
 1039 NSLDLGLPSPNAGQSDERALLDQLHTFLSNTDATGLEGIEDRALGIDELVNOQO--ALE 1096
 1245 LASNLMESEVVRVALTILTKQPVELLQTSNLGICIKG--NCELPNHFPSIMMLNI 1303
 1097 SKQVFOQOEAIVM-----DQKALYQOTYPAQGPPLQGGFNLDGQSPFSNMGCQI--- 1149
 1304 LEAHTDVYAGLHVLSKIMSNKMRHMMHLELILKIKCYOVSKEALRDISMIPR 1363
 1150 --SQGSPFLQGMTPRAGVPR---KINTPKQLAMOLQORLQGGQFLNQSQALEMKEN 1204
 1364 IAPSLPLDLSINIVPIANGEPPTNLCAKILLEVTBHH 1403
 1205 PAGTA-----VVRPMMPQAFVNAQMAA--QOKKELSMHH 1236

RESULT 5
 US-11-055-035-1
 ; Sequence 1, Application US/11055035
 ; Publication No. US20050256072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: AKONIN, NEIL
 ; APPLICANT: ZAMORE, PHILLIP D.
 ; APPLICANT: BRODERICK, JENNIFER
 ; TITLE OF INVENTION: DUAL FUNCTIONAL OLIGONUCLEOTIDES FOR USE IN REPRESSING
 ; FILE REFERENCE: IMY-095
 ; CURRENT APPLICATION NUMBER: US/11/055, 035
 ; CURRENT FILING DATE: 2005-02-09
 ; PRIOR APPLICATION NUMBER: 60/543,467
 ; PRIOR FILING DATE: 2004-02-09
 ; NUMBER OF SEQ ID NOS: 107
 ; SOFTWARE: Patentin Ver. 3.3
 ; SEQ ID NO 1
 ; LENGTH: 3144
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-11-055-035-1

Query Match 1.9%; Score 146.5; DB 7; Length 3144;
 Best Local Similarity 17.7%; Pred. No. 0.089;
 Matches 289; Conservative 213; Mismatches 527; Indels 605; Gaps 71;

193 VEIKVHGDRLRPLRMDVPAKSL--AMLE-----OKFQVQOEGLLPSALRNT 242
 131 MELFLCSDAESVVRKVADECKNKVTKALMDSNLPRQLLELYEIKKNG--APRSRLR-- 186

QY 243 NNGVGLDEADNIGLERPTMIKRP-----LHSAVSSSLRPKNVNDVTGAGAVTMS 297
 DB 187 ---AALMFPAELALVBPQK--RPYLNNLLPCLTRTSKRPDESVOETLAADVXIMAS 240
 QY 298 FESSFEVVPQNLIFAKMDDIYKQVLVIISDKNMDKERVDAKKI---RALLLSYHT 354
 DB 241 FGN-----FAND-----NEIKVLLKAFIALKSSSPTIRTAAGSAVSIQSHR 284
 QY 355 QGFVAVOLKELSLFVDILKEELRSQVIREACITAYMSKTLRNKLDFAFCSLIEHLIN 414
 DB 285 RQYVYVSMNLNLGLLVEDEHSTLLGLVLTIRY-----LVP 325
 QY 415 LIQNSAK--VIASASTIALKYIIKYTHAPKLLKIYDTLNOQSKXDIRSTLCEMLVLF 472
 DB 326 LIQQVQKOTSLKGSFCVTRKEMEVSPPASQVQVELTLHHTOHQ-----H 372
 QY 473 EMQTALENNATVLRDTLKKSIGDADCDARRSRVAYAFRRHFPRLAQIYGTIDIAQ 532
 DB 373 NVVTALF---LLOQ-----LFRTPPELLOTLTAVAGIGQL 406
 QY 533 PALEREREGGGGGT-----GNGTGTAP-----ETRTVSR 564
 DB 407 TAAKEBSGGRSGSIVELIAGGSSCSVLSRKQKXVLLGBEBALDDSESRSVSS 466
 QY 565 GRTPGTLQKPTPSMRSISAVDTAAQRAKV---RAQYTLVSROKPLGPNNSNOAS-- 617
 DB 467 ALTAIVKOBISGBELAASSGVSTPGSAGHDIIITEQPSRQHTLOA-----DSVLASCD 518
 QY 618 MTGAASG-----SLPR--PRLNSNGGTPTATPGSVTPRRRGAAGVSQSP 662
 DB 519 LTSSATDGDDEEDILSHSSQVSAVPSDPAMDND--GTOASSPSIDSSQ-----TTTEG 570
 QY 663 GSRSTSPSTKLDDQXGIGNYRG-----ATGAIPKASGIRSTA----- 703
 DB 571 PDSAVTPDSSEIVLDGTNOYLGLOIGQPODEDEATGILPEASEARNSMALQQA 630
 QY 704 -----SSRETSPTRSSGG-----LM 718
 DB 631 LKKNMHCQPSDSSVDKFLVDEATEPGQENKPCRIGDIGQSTDDSDAPLVHCVRLL 690
 QY 719 KSMVSTGSGSRTPERRNNPVRPSAARLLA-----Q 750
 DB 691 SASFLITGKXVLPDRD--VRVSVKALALSCVGAVALHPESFESKLYKVPDLTTYPE 748
 QY 751 SREAEHTLGVDGDDGPD-----VYSGDYMSGGMR----- 780
 DB 749 EGYVSDILNYIDHGDQVRGATAILCGTLICSLISRSRHHVDDM--GITRILTGTNTSL 806
 QY 781 ---MGRKLMGRDESD-----DIDSEASSVCSESRFDSSTYTRGNKSNYSLSGSHR 827
 DB 807 ADCIPILRKTL--KDESSVYCKLACTAVRNCVMSLCS-----SSYSELGLQI 852
 QY 828 LDMSTORAPFDIETIIQFCASTHSEKRD----- 857
 DB 853 IDVLTLR-----NSSYMLVTRTELETTAEIDFRLVSFLAANAENLRGAHHY 899
 QY 858 -GLISLTQYLAAD-----GKE-----LTOQQLKCVLDMFPRKMDHTTKYSLFLD- 901
 DB 900 TGLLAKQERVLNNVILHLGDEDPVRVHVAASLILVPLKFLYKCCOGQADPVVAARQD 959
 QY 902 -TYTELLIVANETSHNGSSSCLTRLFNKLTGDLNS-----MSHKIMTLQV--HEVPP 954
 DB 960 SSVYKLMLHETQPRPHFSVSTIRIYR--GYNLLPSIDVTMENNLSRIAVVSHELIT 1017
 QY 955 TOLQ-----LKELFRII-----SDSTQPTTKTRAIIRFLT 986
 DB 1018 STRPALTFGCCALCLSTAFPCVCSLMSHMGCVPLSASDESKSCVGMATMILLIS 1077
 QY 987 DLANTYCKSSDFSDOSQACERIVLKLQALQADQKSELRQSARSCVLVNLNTPQMTL 1046
 DB 1078 ---SAMPFLDLS-AHQDALILAGNLLAASAPKSLNSVASSEEA-----NPAATK 1123

QY 1047 -----LLAD--LPRKYQDSARSCISHMRQSSC-----NSGANS 1080
 DB 1124 QEEVWPAIGDALVPMWEQ-----LPSHLKVINICAHVLDVAPGPAIKALPSTLNP 1177
 QY 1081 FSSSPL-----SSSSPKPLOSPSVGPASLQSHHQLSISSTSPRSOSVE---- 1127
 DB 1178 PLSPIRRKKGKEKEGEGASVPL--SPKGSASMAAS--RQSDTSGPVTTSSKSSLSGSFYH 1234
 QY 1128 -----OELL-----FSELDIQHNIQKTSBEIRHCFGQOYQALAPNGENGLQYHD 1174
 DB 1235 LPSYLKLDVLAATHANKVTLIDONSTEX-----FGGFLRSAL----- 1273
 QY 1175 QGGQDSCASLSNKGTOSSANTQO-----SNTPESATMRDNTERERTONAKSP 1224
 DB 1274 ---DVLSQIIELATLQDQIGKCEVEILGYLKSCTSRPMATVVCQQLKTLFGNLSAQ 1329
 QY 1225 TDDAKVITVSINMAENGELILASINMESEVVRVALTLTKDQPVELLQSLTNLGI--CI 1281
 DB 1330 FDG-----LSNPSKSQ-----GRAQRLGSSSVRPGLYHYCF 1361
 QY 1282 KQNGCELPKPRSIM--RMLNLTIEAHTDVVIALGLVLSKI--MRSN-----KMR-- 1329
 DB 1362 MA-----PYTHFTQALDASLRNMVQAEQENDTSGMPVLQKVSQTLKTNLTSTKNRAD 1416
 QY 1330 ---HNMHPLELILKTIQ-----CYQHSKALRDI DSMIPRIAPSLPLDSINIYVP 1379
 DB 1417 KNAIHNRILPEPLIKALKQYTTTTCVOLQOVLDLALQVLQVRNYCLDSDQVFIG 1476
 QY 1380 V-----IATGEFP-----TNLCAIKILEVTEHGESEITDMLDIVFPNLARSADDTQ 1427
 DB 1477 VLKQREYIEVQGFRESEAIININIFVLVLSYERHSQI-----IGPKIQLQCDGIM 1530
 QY 1428 SMVRKAAVFCIVKL 1441
 DB 1531 ASGRKAVTHAIPAL 1544

RESULT 6

US-10-793-626-2964
 ; Sequence 2964, Application US/10793626
 ; Publication No. US20050255478A1
 GENERAL INFORMATION:
 ; APPLICANT: KIMBERLY, WILLIAM JOHN
 ; TITLE OF INVENTION: STRAPHLOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
 ; FILE REFERENCE: PUS4800S
 ; CURRENT APPLICATION NUMBER: US/10/793,626
 ; CURRENT FILING DATE: 2004-03-04
 ; PRIOR APPLICATION NUMBER: 60/164,258
 ; PRIOR FILING DATE: 1999-11-09
 ; NUMBER OF SEQ ID NOS: 4472
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2964
 ; LENGTH: 5024
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic
 ; OTHER INFORMATION: amino acid sequence
 ; FEATURE:
 ; NAME/KEY: MOD RES
 ; LOCATION: (5024)
 ; OTHER INFORMATION: variable amino acid
 US-10-793-626-2964

Query Match

1.9%; Score 146; DB 1; Length 5024;

Best Local Similarity 17.9%; Pred. 0.19; Matches 219; Conservative 233; Mismatches 590; Indels 456; Gaps 66;

QY 88 TATVLPVHYDR--LQSDRDYREKAOQLLRDLMRHVRVLPALIDKALTSQFKKQNAV 144
 DB 2876 TRSEVHOYVINKAQLNDSMTLRQS-----ITD--SHEVQTSNYINETVGNQTAAYNAVD 2929
 QY 145 REEFLQTIYNALHEGTQOLSVRVYIIPVYCALLGDPYVVRBAQTLVEIKYHVDRLR 204

```

Db      | 2930 R---VQOIIHQ-----TSNPTNMPLE--VERATSNVTKSKDALKH 2963
Qy      | 205 PDLRMDVDPASLAME--QKFDQVQOGL-----LLPSALKRTNGVGLD----- 250
Db      | 2964 GE-RELNDKNSKTFVAVHLDNLNOAQKALTHEIQOATIVSOVNNIYKAKALINNDMK 3022
Qy      | 251 -----EADNIGLREPRPMIKRPLHSVAVSSSLRPKNVDYTGAGAVTMESFESSFEV 304
Db      | 3023 LKDIIVAQOQNV-----RQSNYINEDSTPQNMVND-----TIMHAQSIIDQ 3063
Qy      | 305 VQOLINFAKMDDDIYKQVLVITISQKA--DWEKQVADLKKIRALLILSH--TOPQFVAV 361
Db      | 3064 VANPTMSH---DEIENAINNIIKHAJNALDGEHKLOAKENANLLINSJNDJNAPORAI 3119
Qy      | 362 QKELSLSEFVILKEELRS--QVIREACITIAVMSKTLRNKLDACFWSILEHILNIONSA 420
Db      | 3120 NBLVNEAQREKVAEBLOSAQALNDA-----MKHLRNSIQONS 3157
Qy      | 421 KVIASASTIALKYI-----IKYTHA-----PKLLKIYDTLNOQSKS- 456
Db      | 3158 SVRQES-----KYINASDAKKEQYNHAREVENIINEQHPTLDKEIHKQLTDAVNOAND 3212
Qy      | 457 -----KDISTCEIMVLFEEMQTKALER--NATVLRDTLKSIGDADCDAR 503
Db      | 3213 LNVGELLDADKONAHOSIPTLMHL--NOAQONALEKINNNAVTRAKVAALIQOAK----- 3265
Qy      | 504 HSRVAYMAFRHFPBLADQIYGTLDIAQALEREEREGGGGGGTGTGTAPETRRT--- 560
Db      | 3266 -----ILDHAMNLEESIKDKQOVQOSSNYINEDVDQETVYN 3303
Qy      | 561 -VSIRIGRPTGLQKPTPMRSIS-AVDTPAAQRAKAVAYTLY-----SRQKRLGPNN 612
Db      | 3304 AVDHTEILNQTVNPSTIEDIEHAINVQAKKOLRGKOKLYOTDILADKELSKLDLTL 3363
Qy      | 613 SNOQMTGAAAGSLPRPLNSNSGCTPATTGCVTPRRRAGVQSGQSPGSRSPSK 672
Db      | 3364 SOOSSSI-----SNOIYTAKT-----RTEVAQALEKKSINHAMK 3398
Qy      | 673 LRDQVGGIGNYRGATGAIPIKASGIPRSTASSRETSPTRSSGGLMKSMTSGASRRT 732
Db      | 3399 -----ALNKIYKADKVLDSRINEDQPEKA---YQQAINHVDIIHR- 3440
Qy      | 733 PERNNPVPSPAPARLLAQSREAEHTLVGDDQDPYVSGDYNSGGMGRKXL--MGRDE 790
Db      | 3441 -QTNEMPTVINSTIHELETAQNNUL-----HGD-----QCLAAHAKQDA 3478
Qy      | 791 SDDIDSEASVCSER---SFDSSYTRGKNSYVSGSHTRLDMDSTORAPFDITIIQF 846
Db      | 3479 ANVINGLJHLNVAQREVMINTNTNATTRKVKAKNDNAQA--LDKA-----METLOQV 3529
Qy      | 847 CASTMSEKDLISLTQYADGKELTQOQKCVLDMFKMMDHTKYVSLFLTVLVEL 906
Db      | 3530 VA-----HKNNIILDSKILNEDSKYQOQYDRAVIDAEBOLLQTNPTLEPKYIVIVDN 3583
Qy      | 907 ILVHANETSRRNGSSCLTRLFNKLGTDLLNS---MHSKIMWTLQ-----VHAEYEP 955
Db      | 3584 VL--ANEKILFGA-----EKLSDYKSNANDEIKHMMYLNNAQKOSIKMISHAALRT 3633
Qy      | 956 QLO-----LKEPFIISDSTQTPPTK-----RIALRFLTDLA 989
Db      | 3634 EYKOLLQOAKTIDEMKSLIEDTQVAVITDTLLPNTYASEDKKEXVDQVSAQAIIIXI 3693
Qy      | 990 NTYCKSPFPSPDO--SOACERTVLKLAQLAADQSMELRSQARSCLVALYNLMTPOMTLL 1048
Db      | 3694 N---GSNVSLDQVQALEQULTQASENLDDGRVEAKYHANOITDOLTHLNSLO----- 3744
Qy      | 1049 ADLPRVYDSARSCJHSHMRQSGSCNGANSPPSSSPKPLQSPSVGPFASLOSH 1108
Db      | 3745 -----QOTAKESV-----KNATKLEIATATSNMNLANKV--MOKLBOF 3781
Qy      | 1109 HHQJLSISTSPRSRSSVQOELLFSSSELDIGNIK--TSEIHHCQGOVQYTLAENG 1165

```

```

Db      | 3782 INHADISNDNYROADDKIIAYDADLEHGODIQKSNATONEAKQALQULINETSJNG 3841
Qy      | 1166 FNGLQYHDQOQDSCASLSNSKTOSSA---NTTOSN----- 1200
Db      | 3842 FE-RUNHARPALEIYISLEKINNAQSAALBDKYQSHDLLEHLVNEGTLNDIMGEL 3900
Qy      | 1201 -----TESAT---WRLDNLEBERTO---NAKSTFDKAYTIVSINM----- 1237
Db      | 3901 ANAIVNVAFPKASINYINADNLKKNFTQAINARADALNTQOGNIDFNALIDPFKDI 3960
Qy      | 1238 -----AENG-ELIASNMESEVVALTLTDOFP---ELLQTS-----LTLN 1277
Db      | 3961 KTKOALNIEKLTAKSAFAELTIDSLKFTINAOFTHANDEIMNTISIAQLSRIYQADL 4020
Qy      | 1278 GICIGGNCGLPNHFRSINRMLNITLAEHTDVVIAGLAVLSKIMSKRKHMMHFE 1337
Db      | 4021 NDAMKSLRDELNLNQAF--VQASSNYINSDE-DLKQGFDAHSN-ARKVLAKENGNLDE 4076
Qy      | 1338 LILKTIQCYQSHKDALDIDSMIPRIAPSLPDLISINIVPV-----IATGEFTN 1389
Db      | 4077 IQIBGLKQVIDTDALNGIQRLSKAKAKAIQYVQSLSYINDAQRHIAESNIHNSDDL 4136
Qy      | 1390 LC-----AIKILEVTEHHSSEI-----TDHLDIVFVPLARSADDTOS 1428
Db      | 4137 LANTLSKASDLDMKMDRLDTLESNSTSVNSVNYINADKRLQIEFDEALQOASATSS 4194

RESULT 7
US-10-131-826A-10
; Sequence 10, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroft, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guthrie, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria A.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588

```

PRIOR FILING DATE: 1997-09-19
 Remaining Prior Application data removed - See file wrapper or PALM.
 NUMBER OF SEQ ID NOS: 550
 SEQ ID NO 10
 LENGTH: 594
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-131-826A-10

Query Match 1.7%; Score 133; DB 1; Length 594;
 Best Local Similarity 23.6%; Pred. No. 0.063; Indels 116; Gaps 24;
 Matches 118; Conservative 55; Mismatches 210;

443 LKITYTDLN---QSKKDRISTLCELMWL---FE-----EMQTKALBRN 482
 133 LYQIYIDLYGGLQRPSEDEKKLAEKASIGYTIEDSTVAEVEKAEPSEESAADEE 192
 483 ATVLADTLKKSIG-DADCDARRHRYAYWAFRRHPELADQ--TYGTLD----- 528
 193 SNSDEDEVIPDIVADVDELNOEQA-----DLNKQATYGMADGDGVFVRLRKDK 243
 529 -----IAORALEREDEGGGGGTGTGTAPETRTVSRIGTPTGTLQKPTPSMRISA 583
 244 EEAERIKIAKALEEKAMYSGRSRRRQRRREFREKLRGKIS-PSYARROSPTIDPYKR 302
 584 VDTAAQRAKYAQYTLYSRQKPL-----GPNNSQASMTGAASG-SLPRPLNSNS 636
 303 SPSESSSRSRSRSPPTGRBEKIFITISFGSDDEAAAAAASGVTTCCKPAPAPQP 362
 637 GGTPTATPG-SVTPPRRAGVSGSQ---PGSRSTSPSTKLADQYIGYIGYRGATGAI 691
 363 GG-PA--DGRNASARRSSSSSSSSASRTSSSSSSSSSS--RSRRG--GGYRSGRHAR 416
 632 PKKASGIRSTRSRETPTREGGGLMKRSMVSTGAGR-----RTPERNNPVAPSPA 745
 417 SRSRW-SRSRSRSHRRYSRSR-----RGRHSGSGSDGHRYSRSPRRGCGPR--- 466
 746 RLAAQRAEHTLVGDGQPDYVSGDYKRSQGMGRKLMGRDESDDIDSEASSVCSE 805
 467 ---RSTRSRSH-----SGDRYRRG---GKLRHSSSRSSSSLSLPSRSR 506
 806 SPDSYTRGNKSNVLSGSHRTLDMSTORAPPDIEITIIQFCASHWSE-----R 855
 507 SLTRSRSHSPSPSGRSRSR---SQSPSPAPAREKLTTPAASPAVGEKLTKEPPAAGK 563
 856 KDGLSLTQYLADEKELTQ 874
 564 ETGAAKVTQADASGEAETE 582

RESULT 8
 US-10-821-234-972
 Sequence 972, Application US/10821234
 Publication No. US20050255114A1
 GENERAL INFORMATION:
 APPLICANT: Labat, Ivan
 APPLICANT: Stache-Crain, Birgit
 APPLICANT: Andarmani, Susan
 APPLICANT: Tang, Y. Tom
 TITLE OF INVENTION: Methods for Diagnosis and Treatment of Presclampsia
 FILE REFERENCE: 821A
 CURRENT APPLICATION NUMBER: US/10/821,234
 PRIOR FILING DATE: 2004-04-07
 PRIOR APPLICATION NUMBER: US 60/462,047
 NUMBER OF SEQ ID NOS: 1704
 SOFTWARE: pc_seq_genes Version 1.0
 SEQ ID NO 972
 LENGTH: 482
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-821-234-972

Query Match 1.7%; Score 131.5; DB 1; Length 482;
 Best Local Similarity 20.1%; Pred. No. 0.059; Indels 255; Gaps 34;
 Matches 127; Conservative 67; Mismatches 162;

486 LRDTLKKSIGDADCDARRHRYAYWAFRRHPELADQYGTLDIAAQRALEREDEGGGG 545
 7 LRD--RRGLDQOATTPARRELLSALPHAPR-----QARWEPAASSGGGG 51
 546 GTGTGTGTAPETRTVSRIGTPTGTLQKPTPSMRISAADTAAQRAKYAYTLYSRQR 605
 52 GGGGGWHGP-----GRCP-----PCGDDEDTAAMDSPC-----QP 82
 606 KPLPNNNSQASMTGAASGLPRPLNSNGGTPATTPGVTTRP-RGRAGVSGSQSGS 664
 83 QPL-----SQALPQ-----LPGSSSEPLEPPEPGARNGVESLYLP-- 116
 665 RSTSPSTKLADQYIGYIGYRGATGALPKKA--SGIPRSTASRETPTPSG-----CG 716
 117 -----CPLPSTYRCP-----VPSEASAGSCTPRATRTTASPLADFGCGDQGG 161
 717 LMKRSMYSTGA-----GSRRTPRRNNPVRESAPARLLA-----QGRE 753
 162 EL-RPLQSEGAALVTGCGCLAAQG--ARPEAPKRWAEADGDGAPSPKAPWAOENQE 218
 754 AEH-----TLQVGD-----DGQPDYSGDY---NRSGWRMRKLMGRDES 791
 219 AEREGMCSGSCSGSGEASAGLMEALPSAPERLALDVIYPCMYGICVQDSFLGALG 278
 792 DDISEASS-----VCSERSFDSYTRGNKSNVLSGSHRTLDMSTORAP----- 836
 279 GRVLAEEVALKRGRLADQGLVSGRAIPRSRKD-----QLAWEGHPEGGRS 327
 837 ---FDIEITIIQFCASHWSEKXGGLSLTQYLADEKELTQOOLKCV---LDMFRMF 888
 328 IGAAMAHVDAIYRHCAG-----RLGSYINGR--TKAWACYPGNGLGVRRHV- 373
 889 MDHTKYVSLPLDTVTLEILLVANETSRNSSCLTRFNKLGTDLNL---SNHSHIW 943
 374 -----DNPHGGRGICITCY-----YLNQNDVYKHGGL- 401
 944 KTLGVHEVPTQOLKELF-RII---SDSTQ---PTTKTRIALRLFTLDLANTYCKS 995
 402 -LQTFEGRPVANIETPLFRLLIFMSDRNNPHEVKAAYATRAIITWYTD----- 451
 996 SDPDSQOACERTVLK-LAQLAADOKSMEL 1025
 452 -----AKERRAAKDKYQLASGQKGVQV 473

RESULT 9
 US-10-971-982-3
 Sequence 3, Application US/10971982
 Publication No. US20050244889A1
 GENERAL INFORMATION:
 APPLICANT: Rosenfeld, Michael G.
 APPLICANT: Glass, Christopher K.
 APPLICANT: Rose, David W.
 APPLICANT: Torchia, Joseph
 TITLE OF INVENTION: A Transcription Factor Coactivator Protein, p/CIP
 FILE REFERENCE: 6627-PAL021
 CURRENT APPLICATION NUMBER: US/10/971,982
 PRIOR FILING DATE: 2004-10-21
 PRIOR APPLICATION NUMBER: US/09/445,353
 PRIOR FILING DATE: 2000-05-05
 PRIOR APPLICATION NUMBER: PCT/US98/12263
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/049,452
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 3
 LENGTH: 1463
 TYPE: PRT

ORGANISM: Mus musculus
US-10-971-982-3

Query Match 1.6%; Score 125; DB 1; Length 1463;

Best Local Similarity 17.2%; Pred. No. 0.79; Indels 454; Gaps 56;

Matches 240; Conservative 194; Mismatches 511; Indels 454; Gaps 56;

171 PFCVALLGPTVNVREAAIQTIVEIKYGVDRLPRLRMDDVPASKLAMEQKPDQVQ 230
172 PFCVALLGPTVNVREAAIQTIVEIKYGVDRLPRLRMDDVPASKLAMEQKPDQVQ 230
67 POKALIKETVKQIRQ-----IKGEKAAANIDEVQSDVSTQGVLDKDA 114
231 EQLLPALAK-----NTNGVGLDEADNIGREPTMTIKRPL-----HSAVSS 276
115 LGRPMLEALDGFEEVNVLEGSVVFERNVTOYLRYVGEIMNKSVYSILHVDHDEFVKN 174
277 LARKPNVN--DVTGDAVATMESFESEFEVPOQLNFIHAKMDDIYKQVLIIISPKNDM 334
175 LIPKSNVNGSWSGEPPRRSSHTFNCRLVKP-----LPDSEBG 214
335 EKRVDAKKIRALLILSYHTOPFVAVOLKEISLSEV---DILKEELRQVIREAC-- 387
215 HDSQEAHQYEAHQCAV--SQPKSIKEGEDDLOSCLIVHEDPHBEKTNSSILIRKLYHP 273
388 -----ITAYMSKTLRNKLDACFWSILEHLINLIQNSAKYIASASTALK--YIIXY 437
274 GPPRODHFTGHVHNS-----RHEAGLGRSGKDAFRSSTHMSKGLYHMPR 320
438 THAPKILK-----IYDTLN-----OSKXDIRSLC--ELMVILFEEMQTKALE 480
321 RHHEVLRGGLAFSQTIRPSLSDGTLVAAQTKSLRSTOTNEPOLVLSL---HMLHRE 376
481 RNATVLRDLTKKSGDADCDARRHSRYAWAFRRHPELADQIYGLDI---AAORAL 535
377 QNVCVANPDLTGQAMGKPLNPISSSSPAHQALCSGPDQDMLSSININPMNGPKRQGM 436
536 EBERGGGGGGTGTGTGTGTAETRTYRISIGRTPTGLQKTPSMRSISAVDTAAQRAKVR 595
437 PWGRFGSGG-----MNVSGQATTPQ---G 460
596 AQTLYSRORPKPLGPNNSQAQMTGAASGLPRPLNSGSGTPTATPPGSGTTP----- 649
461 SNYAL-----KXNSPSQSPGMNPGQASSVLSPRQMSGVAGSPRIIPSGSPGNLHS 515
650 -----RPRGAVSOGSPGSRSTSPST 671
516 PUGVCSSTGNSHTNSSLNALQALSEHGVSIGSSLASPDLKMGNLQNSPVNMMPPLS 575
672 KL-----RDQVGGIGNYTRGATGAI-----PKKASGIPRSTASSTETSPTSGGLMKRS 721
576 KMGSLDKXDCFGLYGEPSPKGTGOAEASCHPKKOKG-PNDSMPQAASGDRAEG---HS 630
722 MYSTAGSRRTPE-----RNNVPRPS-APARLLAQRREAHNLGVDDQDPDVSDVMS 776
631 RLHDSKGQTKLQLLITTSDDQEPSPPLPSLSDTNDS-----TSLPQPSGTHGTSLE 685
777 GGMWGRKLMGRDESDID--SEASSVCSERSFPSSYTRGNKNSVLSGSHTRLDWSTQ 833
666 KHKILHRL--QDSSSPVDLAKLTAEATKELSSQSSST----- 722
834 RAPFDIETIIOFCASHTWSEKRDGLISLTOYLADQKELTQOOLCVLDMFRKMFMT-- 891
723 -APGSEV-TVKQEPASPKEEN-----ALRYLLKXDKDIDKGLPEITPKLERLDSKTD 775
892 -HTVYVSLFLDTVELLIVANETSRRNSSCLRLFN--FLGDLNLSHMSKIMKTLQ 947
776 ASNTKL--IAKTIVE-----EVSFEPSDQPSGLDLELIDDLQNSQ----- 817
948 VVHEVFP-----TOLKELFRILISDTOTPTTKRIAL----- 982
DB -LPQFPDRPGAPTPGSVVKQAIINDLMQLTADSSVPVPAQAQKALCMSSGFFNNPRG 876
QY 983 ---RL----- 985

DB 877 QLGRLPYQNLPLDITLQSPGAGFPPIRNSPSYSVIPQGMNGNQMILSGNLGNS 936
QY 986 TDLANTYCKSSDPPS-----DOSACERTVLLKLAQADQKSMELRSQARCLVALYNINT 1041
DB 937 TCMIGSSTRSRMSBGEWAPOSTSCESTVLLPLVPRDQKEAR-----RCN 984
QY 1042 PQTLLADLPKIVQDSARSCISHMRQOSCNSCAN-----SPSSPLSSSS 1090
DB 985 P-----TASIM-----GANSOLQORQMQLQSVNMIIGSELEMMNGPQYNOQAAPPNOTA 1035
QY 1091 PKPLSPSPVGPF--ASLOSHHOLISS-----TSPRASQSVQ----- 1128
DB 1036 FWP---ESILPIDQASFSQNRQPPGSSPDLLCPHPAESPSDEGALLDQLYALRND 1092
QY 1129 -----ELLFSE-LDI--QHNQKTS---EEIRHCFGQY--QTALAPNGFN-- 1167
DB 1093 GLEEDRALGLPELVSGQADADQFSSQESSIMLEQKRPVPPQYASQAQAGQYPM 1152
QY 1168 GHOYHDQGDQSCASLSSSKTOSANTTOSNTPEATMELDN--LEBERTONAKSPTD 1226
DB 1153 QDPNHTWGRPNYTTILMQRPGLRPTGIYQONPNQLQLQHLQAOQRQPLMNOIS 1212
QY 1227 DAKYITYSIN-----MAENGELLASNMESV---VRVALTLTKQPVLL 1270
DB 1213 SVSNVNTLRGVPDQAFINQMLAORREILNHLRQROVOOQVORTLWRRGGLVNT 1272
QY 1271 QTSNLNIGICIKGNCCLP 1289
DB 1273 PSMVAPAGLPAAMGNPRIP 1291

RESULT 10
US-10-984-645-2
Sequence 2, Application US/10984645
Publication No. US20050244386A1
GENERAL INFORMATION:
APPLICANT: Habener, Joel
APPLICANT: Zulewski, Hendrik
APPLICANT: Abraham, Elizabeth
APPLICANT: Vallejo, Mario
TITLE OF INVENTION: METHOD OF TRANSPLANTING IN A MAMMAL AND TREATING DIABETES MELLII
TITLE OF INVENTION: A NESTIN-POSITIVE PANCREATIC STEM CELL
FILE REFERENCE: 3284/1223
CURRENT APPLICATION NUMBER: US/10/984, 645
CURRENT FILING DATE: 2004-11-09
PRIOR APPLICATION NUMBER: US 09/731,255
PRIOR FILING DATE: 2000-12-06
PRIOR APPLICATION NUMBER: US 60/169,082
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: US 60/215,109
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: US 60/239,880
PRIOR FILING DATE: 2000-10-06
NUMBER OF SEQ ID NOS: 55
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 1618
TYPE: PRT
ORGANISM: Homo sapiens
US-10-984-645-2

Query Match 1.6%; Score 124; DB 1; Length 1618;
Best Local Similarity 19.5%; Pred. No. 1.1;
Matches 163; Conservative 93; Mismatches 331; Indels 250; Gaps 34;

QY 526 TLDIAAORAL--EREBGGGGGGTGTGTC-----TAPETRTYRIS----- 564
DB 295 SLEAVATRYLLLEANSRLQTPGGGSKTSLSPDQPKLELOPFRTPBGRRLGILBLVSPS 354
QY 565 -GRTPGLQKTPSMSISAVDTAAQRAKVAQYTLVSR-----GRKPLGPNNS 613
DB 355 LPSPLPATLTPV-----AFUKQGEFLQARTPTPLASTPIPTPOABSPA 399

```

QY 614 NOASHTGAAGSLPRPLNLSNGGTP-----ATTGSGVTPRPRGRAGVSGOSGSGSST 667
DB 400 VDAEIRADAPLISLQTOGRKQABEPRAEARVAIPASVLPGP-----EEBQ----- 447
QY 668 SPSTLRQOYGGIGNYRGATGAIIPKASGI--PRS--TASSRETSPTSGGGLMKRSMYS 724
DB 448 -----GGRQASSTGQSPEDHSLAPLSPHSSLEAODGSGGSRVSTCRG 494
QY 725 TGAGS--RRTPERNNVPRSPAPARLLAQSREAEHTLVGDGQPDVYSGDYMRSGEMRGR 783
DB 495 EGEQIIMWLVEKETIIEGVVSSLOQEIWEED-----LNK 530
QY 784 KLMGRDESDIDSEASVCSERSFDSYTRGNKNSYLSGSHTRLDWSTORAP--PDDI 840
DB 531 KEI--QDSQVPLEKETILKSIGEIEIGSLKLENO-----SHETLERENQECPRSLERDL 582
QY 841 ETI-----IOPCASTHWEKDGILSTQVLAQKELTO-----QOLKCVL 881
DB 583 ETLKSLERENKRAIKGCGSSETSRRK--GGRQKLPFGKEDTQTLQSLQKQENQELMKSL 638
QY 882 DMFRKMFMDHTKYVSLF-----LDVTELIVHANETSNN--GSSSCLTRLENKLG 932
DB 639 EGNLETFLEPGTENELVSSLOENLESITALEKEN--QEPRLSPREVQDEAL--RPLTKENQ 696
QY 933 DLNSMWSKIMKTLOVHEHYEFTQQLKELEFRISDSTQPTTKRIALRPLTDLANTY 992
DB 697 EPLRLSED-----ENKEAFRLSEKENQEPKLEEDQSIVRPLETEN 739
QY 993 CKSSFPBDSQACERTVYKLAQ-----LAADQSMELNSQASCVLALYNL----- 1039
DB 740 HKSLRSLSEBODETLRTLEKETQORRRSISGEODQTLRPEREKVDLEPLSLQELIARPLE 799
QY 1040 NTPQMTLLADLPKYVODSARSCISHMRROGQSCNGANSPPSSPLSSSPKPLQSP-- 1097
DB 800 NENQFL-----KSLKESVAVAVSLLETILLESLSKAGQENLETKSEPTQAPLMTBEE 853
QY 1098 ---SVGPASLQSHHQLSISTSPRSQSVQE-----LIFSELIDIOHNIQTS 1146
DB 854 INKSGNESRKGNSRTTCVCGSEPRDQTPRGESGILEISGMEPGEFISRGVDKES 913
QY 1147 ---EIRHCFGGQVOTALAPNGFNHLOYHDOGQODSCASLSMSKTOSSANTQSNTP 1202
DB 914 QRLNEEENLKGXEYQESLR-----SLEEQGE-----LP 943
QY 1203 ESATWRL--DNLERRRTQNAKSP-----TDDAKV-----ITVSINMAENGEL 1243
DB 944 QSAVQRMEDTYEKQELQAOESPFGMAGVENDEALINLRQODGFTGKEVEYQEGEL 1000

```

RESULT 11

```

US-11-046-346-1
; Sequence 1, Application US/11046346
; Publication No. US2005025502A1
; GENERAL INFORMATION:
; APPLICANT: Dana Farber Cancer Institute
; APPLICANT: D'Andrea, Alan D
; TITLE OF INVENTION: Method for Determination and Quantification of Radiation or Genot
; TITLE OF INVENTION: Exposure
; FILE REFERENCE: 7032/2082
; CURRENT APPLICATION NUMBER: US/11/046,346
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/540380
; PRIOR FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1451
; TYPE: PR
; ORGANISM: Homo sapiens
US-11-046-346-1

```

Query Match 1.6%; Score 123; DB 7; Length 1451;

Best Local Similarity 17.0%; Pred. No. 1.1; Matches 249; Conservative 212; Mismatches 498; Indels 504; Gaps 63;

```

QY 51 IDGIMPLWTSQHFYIAQKSLSELSBELIKRLGSDNATATYLP--HVIDRLGDSQDTYREK 109
DB 142 IDIQAIAIKTLF--EKLEPEYFE--NRNSDEINFRLLVYSQKMLDRVVDGDLTKYI 196
QY 110 AOL--LRDMENHVL--PPQALIDKLATSCFKHNAKVREEPLOTIVNALHEVTOOLS 165
DB 197 MQLISIPENQHDIITSKFEILDGDS-----QHDVQKELSDLLI-----ENTS 240
QY 166 VRVYIPVCAALLG--DPT--VNVREAAIQTLVEIYKAVGDRLPDLRMDVPA----- 215
DB 241 LTVIIDLVLGSLRDLDPFLKVRQLVMDKLSI-----RLEDIPVLIKFIH 287
QY 216 ---SKLMEQKFDQVQKQGLLPSALKNNTNGVGLDEADNIGLREBRTNIAKRLH 270
DB 288 SVTAMDTELVISELRKIDLOHCVLPRLQASQVTKSKGRASSGNGQSSGQCIILF 347
QY 271 SAVSSLRPKNNVNDVYGDAGAVTMESFESEFEVVPQNLIFHAKMDMDIKQVLVTSK 330
DB 348 DVISALRYEKTISEA----- 363
QY 331 NADWEKVD-----ALKKIRALLSYHTOPQVAVOLKELSLSPVD--ILKEELRSQVIR 384
DB 364 ---WIKALENTASVSEHKVPDLVWL-----FIVSNTQTKYIDRLNKKIRSGCIG 413
QY 385 EACITIAVSKTLNKLDAFCWSILEHILNIONSAKYIASITALKYIIKY--THAPK 442
DB 414 EQLQSTFSVAYL--VTKDMCSITLSLAOSILHSLDOSIISFGSLLYXAKRFPDTYQO 471
QY 443 ---LKKIYDTLQNSKSKDIRSTLCELMVLLFEEMQKALERNATVRLTKKISIGDAD 498
DB 472 EVGALVTHICSGEAEEDDLVYLBELVVL--NPSAMMAHNVFOGIL----- 518
QY 499 CDARRSRVAYMAFRHPELADQIYGTLDIAORALEREBSGGGGGTGTGTAPETR 558
DB 519 ---DYLNDI-----SPQOI 529
QY 559 RTVERIGHTPGTLOKTPPSMSISAVDTAAQAKVRAQVLYLSKORRPLGN--NSNQA 616
DB 530 RKLFFV-----LSTLAFSKONEASHIODMHLVIRKOLSTVFRKYKI 573
QY 617 SMTGAAGSLPRPLNLSNGGTPATTPGVSUTPRPRGAGVSOO-----PG 663
DB 574 GITGAVTMAGI--MAADRESPSLT-----QERANISDECTQVTSILOLVHGS 621
QY 664 SRSTSPSTKLDQYGGIGNYRGATGAIIPKASGI--PRSTASSRETSPTSGGGLMKRSMY 723
DB 622 EQSPQASALYYDEFANLIGHK-----LDPKALBWWGHTICNDFODAFV----- 665
QY 724 STGASRRTPERNNVPRSPAPARLLAQSREAEHTLVGDGQPDVYSGDYMRSG--M 779
DB 666 ---VDSQVPEGDPFF--PVKALYGLEBYDTQDGINILPLFLFSQDFADGGPVTQ 718
QY 780 RMGRKLMGRDSDSDIDSEASVCSERSFDSYTRGNKNSYLSGSHTRLDWSTORAPFD 839
DB 719 ESGGLTV-----SPIC-----LAPY-- 733
QY 840 IETIIOFCASTHWS--ERKDLISLTOYLD--GKELTQOOLKCVLDMFRKMFMDHT 893
DB 734 -FRLLRLCVERQHNQNLBEIDGLDCPIFLTDLBERGKLEMSAKAASFCMSLIF----- 787
QY 894 KYVSLFDTVELLIVHANETSRRNGSSCLTRLPFKLGTDLLNSHSHKIWKLOVHEFY 953
DB 788 ---LTLNMFREIYVAFQOETSPEKNGKVLRLKIVBLQIL-----LEKTLAVTPTDV 837
QY 954 PT---QOLKELEFRISDSTQPTTKRIAL-----LRFYLDLANTYCKSSDFPS 1000
DB 838 PPLGNFDEVTLDI-----TPHTVTAISAKIRKKGIKKQKIDGSKT--SSSDTIS 886
QY 1001 DOSQA-CERTVYKLAQLAAD-----OKSMELRSQASCVLALYNL 1039

```

Db 867 EEKNSCEDPTSHKGLNKEFTGKEKTSLLIANSHPRELDIEVSIHCGVTKIL 946
Qy 1040 NT-----PQMTLLIADLPKYODSARSCISHMRQSGSCGANSPPSSSP 1085
Db 947 DTEHTEATEVQGLPPELLLEBDL-----SOKLEMTLPPPIARR 987
Qy 1086 LSSSPKPLQSPVSPFPASLQSHHQLSISTSPRSROSSVEQELLFSELDIDHNIQKT 1145
Db 988 V-----PFLKNKG-----SRNIGFSLQORS 1008
Qy 1146 SEERHCGGQGYQTALPANGFNGHLO-YHDQGGQDSCASLSSNKTQSSANTTOS-NRPE 1203
Db 1009 AOEIVHC-----VEQLLTP-MCNLEHNIHNTIQ-----CLAENHGVDPGVKVOEYHIMS 1059
Qy 1204 SATKRLDLERERTQNAKSPTDPAKVTVSINNAENGLILASNLMESEVVRVALTLTK 1263
Db 1060 SCYRLLDQIFHGLFAMSGFSGPENQNLISYALH-----VLSRLKQGE-----H 1103
Qy 1264 DQPV-ELLQTS---LTNIGICIGKNGCELPNKHFRSIRMLNLTIE---AETDVVIA 1314
Db 1104 SQPLEELLQSQVHYLQNHQSI PSFOCAL-----YLIRLMVILEKSTASQNKETIAS 1157
Qy 1315 GLHVLKSKIMS-----NKMHNMMHFELELLIKTIQCYHSHKALRDISMIRIAPSL 1368
Db 1158 LARQFLCRWSPSGDKSKSNISNDQHALLCIYL-----EHTESILKALIEIAQVAPEL 1211
Qy 1369 ---PLDLSINIVNPVIATGEFEPT 1388
Db 1212 INSPKDS-----SSTFPT 1225

RESULT 12
US-11-074-176-134
Sequence 134, Application US/11074176
Publication No: US20050250135A1
GENERAL INFORMATION:
APPLICANT: Kluenhammer, Todd R.
APPLICANT: Russell, William M.
APPLICANT: Altermann, Eric
APPLICANT: McAlliffe, Olivia
APPLICANT: Perill, Andrea Azcarate
TITLE OF INVENTION: Nucleic Acid Sequences Encoding
TITLE OF INVENTION: Strees-Related Proteins and Uses Therefore
FILE REFERENCE: 5051-694
CURRENT APPLICATION NUMBER: US/11/074,176
CURRENT FILING DATE: 2005-03-07
PRIOR APPLICATION NUMBER: 60/551,161
PRIOR FILING DATE: 2004-03-08
NUMBER OF SEQ ID NOS: 381
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 134
LENGTH: 1189
TYPE: PRT
ORGANISM: Lactobacillus acidophilus
US-11-074-176-134

Query March 1.6%; Score 122; DB 7; Length 1189;
Best Local Similarity 18.4%; Pred. No. 0.93;
Matches 209; Conservative 194; Mismatches 447; Indels 286; Gaps 52;

Qy 134 TSCFHAKKAKREEFLQ-TYNALHEVGTQQLSV---RVYIPYCALIGDPTVNVREAI 189
Db 81 TLIFDNKRELAFTDQVSVTRRLIRSGDSEFLINNQOVNRMDVTTFLDGISPNLSAI 140
Qy 190 QTLVEIYHVGDRL---RPDLRMDVPAKSLAMLEOKFQDVQKQGLLSPALKNTNGNG 246
Db 141 ISQGRV---DQILNSRPEQRRIITFEAAAGVLFHKQKEAQNQ-----LKTQDNL 188
Qy 247 VGLDEADNIGLERPTMRKRLPLSAVSSSLRPKNVADVTDGAGVMTSESSFEVVP 306
Db 189 IRINDL---VKLESRL--EPLNF--QSLAKEVKFPKSLGDKLKLALFAE-----IE 235
Qy 307 QLN-----IFHAKMDDIYKQVAVIISDKNADWEKRVDAKKIR-ALLILS 351

Db 236 NINQOREIDQKADKNKILAK-LDDEVKDSQAANVQKRAEYKRLRERDHTQKKLKS 294
Qy 352 YHTQPFVAVOLKELSLSPFDILKEELRSQVIREACITIAVMSKTLRNKLDAPFCMSILH 411
Db 295 KDLSELNASTQMAEQSGFDDATKEBYKNQV-----KQLNQ----- 330
Qy 412 LNLIONSAYKAVIASASTIALKYIITHTAPRLKI---YTDILNQ-----SKQDIRS 461
Db 331 -NLVQLKADL-----DELKKEKKLODEODVYKIERGQLTGEINDEPEELNKLDIDRN 383
Qy 462 TLCEIMWLFPEEMQTKALERNATVLRDLTKKSIGDADCDARRHRVYVMAFRHFPFLAD 521
Db 384 NYMQL-----ODQATNNQIV-----NLMSDKRSQ-----AD 412
Qy 522 QIYETDLIA-----ACRALEREREGGGGTGTCTGAPETRRATVRSIRGPTGLQKPTP 576
Db 413 TTYQGVSVSKQLDPAQQLQQLRIEG-----KULTDRKQEQNALIVRINKQNNQLTELT 467
Qy 577 SMSISAVDTAAQAKAVRAQY-TLYSRQKPLG-----PNSNQ-ASMTGAASGS 626
Db 468 NLRQV--VNAERNELEKVEARHEALVNIQKHHEGYVYGVANNVNLHNDPAGVIGAV--GE 523
Qy 627 LPRRLNSNGGTPATPGSVTPRPRAGVSGSQPSGRSTPS--TKLDQYQIGINNY 684
Db 524 LIT-----PPAELEAMTTALGG--GVQDLITERSISARNAINKLKNHGGRAF-- 571
Qy 685 RGATGAP---KXASGIPRSTASRSTSPRSQGGGLMKRSMYSGASRRTPERNNPVP 741
Db 572 ---LPLDGLAQYGIPOSTVTT---LKSVDGF--RGIASDLVBSKTDQDITAAIN 618
Qy 742 SAPRLLAQRERAEHTG-----VGDDQDPYVSGDYMRSGGMGRK--LKGDE 790
Db 619 LLSGVAVITDITDAMSVAQVRNRYRITVLDG--DVLSPGSGMTGGQGNQNSNPLQATE 676
Qy 791 SDDIDSASSVCERSFDS-----YTRGNKSNYSLSGSTRLDMSQRPAPDIEITIQ 845
Db 677 INLEKQIKTKLKNLEDDQKLENLVDQSKVNAELQDLALRETSQ-----AIN 727
Qy 846 FCASTHWSERKDGILSTQYIADGKELTQOOLK---CVLDMFRKMPMDTHTKVYSLPDT 902
Db 728 EALISFGQKE-----VKRLDANTLYKSRIDRNRRIELKKQIEANDK----- 774
Qy 903 VTELILVHNETSRNGSSCLTRL--FNKLGTDILNSMESHKIMTLOYV----- 949
Db 775 --QMLTKQGEQEKAKKNDLQDKIKNFNNLSORIOBEL-SKLPKIAVYTNKLENLSSQE 831
Qy 950 ---HEFFPQQLKEL---FRILSDSTQPTTKTRIAILRFLDLANTYCKSSDFP--- 999
Db 832 NEKNHQIDNSEKQIEDTAVLTLLAODENSMQOT-----ANLEKQKSTIOKN 880
Qy 1000 ---SDQSACERTVTLKLAQADQKSMELRSQAR-----SCVALATNLNTPQ 1043
Db 881 NELQARLNDLSQLGQDQADQINQLDQVAS--RNYDLRKLDAIIEQEDSVIAKRNSSGINQ 938
Qy 1044 MTLIADLPKYODSARSCISHMRQSGSCGANSPPSSSPKPLQSPVGPFA 1103
Db 939 RLEFLRDYSLTFPEAA--IAQAGEENNEETRELNAKSLKLRMSIEDIGFVNLDISIOEYE 996
Qy 1104 SLSQSHHQLSISTSPRSROSSVEQELLFSELDIDHNIQKTSEIRHCGGYQT 1159
Db 997 DVKQDQVFL-----NQOQNDLLFARDLEKSMTELDDEVKTRFPHPTD 1040

RESULT 13
US-10-485-517-141
Sequence 141, Application US/10485517
Publication No: US20050256299A1
GENERAL INFORMATION:
APPLICANT: University of Sheffield
APPLICANT: Biosynexus Incorporated
APPLICANT: Foster, Simon
APPLICANT: Mond, James

TITLE OF INVENTION: Antigenic Polypeptides
 FILE REFERENCE: P100629WO
 CURRENT APPLICATION NUMBER: US/10/485,517
 CURRENT FILING DATE: 2004-02-02
 PRIOR APPLICATION NUMBER: GB 0118825.9
 PRIOR FILING DATE: 2001-08-02
 PRIOR APPLICATION NUMBER: GB 0200349.9
 PRIOR FILING DATE: 2002-01-09
 NUMBER OF SEQ ID NOS: 424
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 141
 LENGTH: 1290
 TYPE: PRT
 ORGANISM: Staphylococcus aureus
 US-10-485-517-141

Query Match 1.6%; Score 121.5; DB 1; Length 1290;
 Best Local Similarity 17.6%; Pred. No. 1.1;

Matches 242; Conservative 186; Mismatches 510; Indels 435; Gaps 56;

51 IDGLMPWLTGSHFKIAQKSLFAFSELIKRLGSDFNATATVLPVHIDRLGDSRDTVREKA 110
 152 IDGL-----SHLTAAQK--EALKQVVOGSTT-----VAEAGN-----182
 111 QLLRLDMEHRYLPPOALIDKLATSCFKHKAQKVEEFLQTVNALHEYGTOQLSVRYI 170
 183 -----EOKANNVDAAMDRLRQSIADNATTKQNTDASQNKKDAYNNAVTTAQGII 234
 171 PPVCALLGDPYNNVREAIQTLVEYKIVGRRLRDLARMDDVPASKLAML-EQCFDQVK 229
 235 DQTSPTLDPYV-INQAAGQ-----VSTTKALNGENLEPAK 271
 230 QEGLLPGLAKNTNG-----NGV-GLDEADNIGLREPTMIRPLHSVAVSSLR 278
 272 QOASQSLGSLDYNNAQKQTYTDQINGATYDEANQIQKQNNLNTANGNLKQALADDA 331
 279 PKPVNDVTGD-----AGAVTSESESEFEV-----POL 308
 332 TKATNPTFDADQAKQAVNTAVTNAENIISKANGNATQAEVEQAIKQVNAKQALANGNA 391
 309 NIFNAKD-----MDDIYQVAVI 326
 392 NVQAKDEBATALINSSNDLNAQKQALQYQVONATVAGVNNVQKQTAELNNAATQKQ 451
 327 ISDK-----NADMEKR--VDALKIRALLISVHTPOQVAVOLKEL--SLSEV 371
 452 IADKEQTYADGNFVNADDPKQVNAVOAKAEAL-----SATPVDVVTPEITLAKNV 506
 372 DILKEELR-----SQVIREACTITAYMSKTLRKLQD-AFCWSILEHLINLIONSAYV 422
 507 TQAKNDLNGNTMLATAKQNVQAHIDQLPRLNAQDEYKQITQATLVVNNVAIQQAATT 566
 423 IASAST-----IALKYIITK-----YTHAPKLIKITY-----DTLN 452
 567 IINDAMTQKQGIANKRAQIKGSENYHDADTKQATVADNAVTKAEELIKQTTNPTMDPNTIQ 626
 453 QSKSDINSTLCEMLWVLFEEWQTKALEENATVLRDLTKSGIDGADCDARHSRYAYNAF 512
 627 QALTK-VNDT-----NQLANGN-----QKADAKQDKK-----653
 513 RHHPFLADQIYGTLD--IAAQRALEREREGGGGTGTGTGTAPESTRVSRIGRTP- 568
 654 -----TTLGTDLHNDQKQALTTQVE-----QAPDI-ATVNNVYKQNO 691
 569 -----GTLQKFTPSMRGSIASVDTAAQAKRAQYITLYSKRQKRLGNNNSQASM 618
 692 NINAMNTNINNALODTETELINSINFID--ADQAKKDAYTNAVSAHEGILSKANGSNAQ 748
 619 T--GAAGGSLPRPLNSGGTPTATPG-----SVTPRPRGRAGVSGSQPG 663
 749 TVEQAMQVRVNEKQALNGNDVQRAKDAKQVITIANDLNQAQMTQKGIADKQOTKAN 808
 664 SRSTPSTLKDQYGGIGNYRGATGAIPIKASGIPRSTASSRETSPT-----RS 713

809 GNFVNADTDKQ-----NAYNNAVVAHEQIISTGPANVDPOQAQALQOVNAQKGLN 861
 714 GGGIMKSMYSTAGSRRPERRNNPVPSAPARLLAOSRAEHTLGVGDGQPPYVSGDY 773
 862 GNNHLQVAKDNANTAIIDQLPNLNOPOKTA-----LKQVSHAEIVTGVNAIKQ-----NADA 913
 774 MRSGMRGKMLGRDE-SDDIDSEASVSCERSFDSSTYRGKNSVSLSGSHRLDMST 832
 914 LNNAMGTLKQOIQANSQVPSVPTQADQKQQAQVYNA-----ANQAOIANGIPI-----964
 833 QRAPFDIETITIOFCASTHWSERDGLISLTOYLADGKELTQOQKLCVLMFRKMFNDTH 892
 965 ---PULTPDVTQ--AVTTMNAQKDL-NDEKLAQAK--QELANLDTLR---DLN 1010
 893 TKVYSFLDVTVELILVHANETSRRSSGCLTRLEN-KLGTDLNSHAKIKWTLOVYHE 951
 1011 QPQRDALRQINQOALATVETQKQNAQNVNTAMSNKQIGIANKDITV-----KASENYHD 1065
 952 -----YPTQQLKELPRIISDSTQPTTTRIALIFLTDLANITYCKSSDPSPDS 1003
 1066 ADADKQATYNNAVSQAGGIINQTTNPTLNPDELTR--ALTQVDAKN-----1110
 1004 QACERTVLKLAQLAADQKSMELRSQARSCLVALYNLNTPOWTLT-----LADLPKV 1054
 1111 -----GLNGEKALTEKONAKDANVSGMTHLNDQKQALQKQIDQSPETATVQV 1159
 1055 YQDSARSCISHMRQSQSCNSGANSPPSSPSSSPKPLQSPSVGPASLQSHHQUST 1114
 1160 KQTN--TSLQAMDQSLAINDKQKQTLADGNYLNADP-----DKQNAVYKQAVA 1205
 1115 SSTSPPRQSSVEBELFSEELDQHNIOKTSERIRHCFGQYOTALAPNGFNHLOYHD 1174
 1206 KAEALLNKQSGTN-----EVOAQVESITNVN-----AKQALNGN-----1241
 1175 QGQDQSCASLSNSKQTSNANTQSNTPESATMELD-----NLERRTTQNAKS 1223
 1242 ---DNLNNAKQAK-QQLANLTLHNDQKQSPSQITQAPLVYDVTITNK 1289

RESULT 14

US-10-499-715-2
 Sequence 2, Application US/10499715
 Publication No. US20050250717A1

GENERAL INFORMATION:

APPLICANT: BENAYAHU, Dafna
 APPLICANT: SHUR, Irit

TITLE OF INVENTION: CHROMATIN REMODELING PROTEIN AS A MARKER EXPRESSED BY STROMAL
 TITLE OF INVENTION: CELLS, DNA ENCODING SAME, ANTIBODIES THEREO, AND METHODS OF

FILE REFERENCE: BENAYAHU-1.1 PCT
 CURRENT APPLICATION NUMBER: US/10/499,715

CURRENT FILING DATE: 2004-06-21
 NUMBER OF SEQ ID NOS: 34

SOFTWARE: PatentIn version 3.2
 SEQ ID NO 2

LENGTH: 2897
 TYPE: PRT

ORGANISM: Homo sapiens
 US-10-499-715-2

Query Match 1.6%; Score 119; DB 1; Length 2897;
 Best Local Similarity 17.9%; Pred. No. 5.3;

Matches 110; Conservative 77; Mismatches 222; Indels 206; Gaps 20;

724 ERNNVPRSPADARLLAOSREAHTLGVDDQCPYVSGMVRSGMRGKMLGRDESD 793
 1693 EKVNTIRADPALCFLEKRGKDEKAVABQANDYMDG-VEDEYKPAIPAIFKDDIED 1751
 794 IDSEASVCSERSFDSSTYRGKNSVSLSGSHRLDMSTQAPFDITITIOFCASTH-- 851
 1752 VSSPGLVYADG--DQLMGMDK-----VWPTQSLATTLRLRLITAYQRTKNA 1798
 852 -----WSEBKD-----857

```

Db      1799  RQIQOIPTEFVSVPNOPIYEATLNPXAKIERQORWTRREAEFYRVVSTFGVFD 1858
Qy      858  ---GLISLTOYLADOK-----ELTQOOLKCVLDMFRKM-----FMDHTKTVYSFL 900
Db      1859  PDRQOPMTKRAARLHKTKDDSLKYLAFMCMCRVCNLPKSELYVD-----NIFI 1913
Qy      901  DVTLELLVHANETSRRNGSSSCLTRLFNKLGTDLLNSMHSKIMKTLOVH-----E 951
Db      1914  QPIRE-----ERASRTLYRIELKRVKQALR--HPQLFRLKLCHEPNPLPWME 1962
Qy      952  YFPFQOLQ-----KELFRITSD-----STQPTTKRTAIALRLDNLANTY--- 992
Db      1963  CGPHDRDLIGAAGKGVSRTHYHLRDELSFMAAQRYSQSKMAHSRTSPLLQYQVA 2022
Qy      993  -----CKSSDPSPDSQACE-----RTVVK----- 1012
Db      2023  LSASPLTSLPRLAKGILLEMVKSENLEKEFQSSSEESMSSEVETRTLIKSEPVAPKN 2082
Qy      1013  --LAQLADOKSMELRSQARSCLVALYNLNTPOMTLTLADLPKYVODSARSCJSHMRQ 1070
Db      2083  GVLPQATGDKSGKGCETDRFMVAARTEPLTPNPA---SKKPRVHKRGSESSSDSDSE 2139
Qy      1071  SQSCNSGANSPPSSPLSSSPKPLQSPSVGFPAQSHHQLSISSTSPRSQSSVQEL 1130
Db      2140  RSSCSRSSSSSSSSCSHSRSRSSSSSSSSSASASS---SSSTSSSSSSSSSSSEE 2195
Qy      1131  LFSSELDIOHNIQKTSSEIRHCFGGQYQTALAPNGFNHLOYHOGGODSCASLNSKT 1190
Db      2196  SDSDEEAQKRETT-----HMKAYD---EBSVASLSTTQDE 2229
Qy      1191  QSSANTQOSNTPESA 1205
Db      2230  TQDSFQMNNGTPESA 2244

RESULT 15
US-10-821-234-1070
; Sequence 1070, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: PL_SEQ_genes version 1.0
; SEQ ID NO 1070
; LENGTH: 1377
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1070

```

```

Query Match      1.5%; Score 116.5; DB 1; Length 1377;
Best Local Similarity 21.3%; Pred. No. 2.7;
Matches 115; Conservative 46; Mismatches 198; Indels 181; Gaps 24;

```

```

Qy      456  SKDIRSTLCMLVLLFEEWOTKALERNATVLRDTLKKSIGDADC-----DARRSHRYAY 509
Db      320  AQNHETILKAKLKEIFRRSKS-----QSSGSAHPISVPGARRH----- 365
Qy      510  WAFRRHPEL-----ADQIYGTLDIAQALEREREKGGGGTGTGTAPET 557
Db      366  ---HHLVNLPPSGTGLVRSRRTSLAAT--PPAAKCSSCRVRTASBGDG--GRAAGAAAA 419
Qy      558  RRTVSRIGR--TPGTLQKPTSMRSISAVDTAAAGRAKVRQYTLYSRQRKPLGPNNSNQ 615

```

```

Db      420  QRLSVAGSEPLSGPVAPRAPSHTLSG--GCCGRGSKVAL-----LPAGALQHS 468
Qy      616  ASMT-----GAAAGSLPRPRLNSNG--GTPATTPGSVTPPR-----GRAGVS 658
Db      469  RSMSPVVAHSPPAATSPGS-----LSSSGHSGSYPPPPGPHPLPHPLHGPQGPSS 523
Qy      659  QSQPGSRSTS--PSTKLRDQYGGIYRRGATGAIPKKAAGIPSTASRRTSPTRSGG-- 716
Db      524  GSASAGSPSPDPGFMGLDEYGS---SPGDLRAFCSHRSNTPESTA---ETPPARQGGG 576
Qy      717  ---LMKRSWYSTGAGSR-----TPERNNPVR----- 740
Db      577  GEFTGYMTMDRPLSHCGRSTRVYSGDAODLRLGRLKRTYSLTPPARQRPVPPSSASLD 636
Qy      741  -----PSAPARL-----LAQSRBAHTLGVGDGQPD 767
Db      637  EYTLMRATFSGSAGRLCPGCPASSPKVAHYHPYEDYDIEIGSHRSSSNLGDADGYPM 696
Qy      768  YVSGDYMRSGGMKMKLMGRDESD-----IDSEASSVCERSF 807
Db      697  TPGAFAFGSGS-----GSCRSDYMEMSPASVAPKQILOPRAAAAAAAVPSAGPA 748
Qy      808  DSGYTERGKNSYSLSGSHTRLDWSTORAPDDIETIIQFCASTHWSERKDG-LISLTOYL 866
Db      749  GPAPTAAAGRTFPASGGYKASSPAESSP--EDSGYKRMCMGSKLSMEHADKLLPNCGYL 807

```

```

Search completed: November 23, 2005, 15:17:09
Job time : 25 secs

```

HEINERSON, L.
101030850 Page 1
Seq IDs 1 & 2

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 14:18:08 : Search time 16694 Seconds
(without alignments)
17518.843 Million cell updates/sec

Title: US-10-030-850-1
Perfect score: 5145
Sequence: 1 gaattcgccagcagcgcgtc.....ggattaatacaataaaaa 5145

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5683141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5145	100.0	5145	AX073964	AX073964 Sequence
2	5120.8	99.5	5959	AB031048	AB031048 Drosophila
3	5055.6	98.3	5916	AF195498	AF195498 Drosophila
4	5055.6	98.3	5938	AF250842	AF250842 Drosophila
5	5055.6	98.3	5962	AY069579	AY069579 Drosophila
6	5051.8	98.2	87197	AC014071	AC014071 Drosophila
7	5051.8	98.2	163098	AC010055	AC010055 Drosophila
8	5051.8	98.2	182960	AC010017	AC010017 Drosophila
9	5051.8	98.2	313157	AB003593	AB003593 Drosophila
10	1932.4	37.4	96091	AC130040	AC130040 Rattus no
11	1834.4	35.7	79733	AC006579	AC006579 Drosophila
12	732.8	14.2	767	AR495636	AR495636 Sequence
13	732.8	14.2	767	AR510918	AR510918 Sequence
14	661	12.8	1420	AR501058	AR501058 Sequence
15	661	12.8	1420	AR516340	AR516340 Sequence
16	452.4	8.8	466	AR509306	AR509306 Sequence
17	331.4	6.4	339	G01355	G01355 fruit fly S
18	149.8	2.9	3153	AX746802	AX746802 Sequence

ALIGNMENTS

19	149.8	2.9	3153	8	AK091140	AK091140 Homo sapi
20	149.8	2.9	7888	6	CQ873990	CQ873990 Sequence
21	149.8	2.9	7888	8	AF347693	AF347693 Homo sapi
22	149.4	2.9	154	6	AR525441	AR525441 Sequence
23	141.8	2.8	2524	6	AX746992	AX746992 Sequence
24	141.8	2.8	2524	8	AK091451	AK091451 Homo sapi
25	140.2	2.7	7426	5	DQ022563	DQ022563 Xenopus l
26	125.6	2.4	4855	5	BC066436	BC066436 Dantio rer
27	116.6	2.3	1583	5	BC029035	BC029035 Homo sapi
28	116.6	2.3	1908	6	AX746588	AX746588 Sequence
29	116.6	2.3	1908	8	AK090716	AK090716 Homo sapi
30	116.6	2.3	5614	6	AX683140	AX683140 Sequence
31	116.6	2.3	5614	8	AB014527	AB014527 Homo sapi
32	116.6	2.3	6445	6	AR338833	AR338833 Sequence
33	116.6	2.3	6445	8	AR338832	AR338832 Sequence
34	112.2	2.2	4494	9	EC039216	EC039216 Mus muscu
35	107.6	2.1	5570	9	RNO276962	RNO276962 Mus muscu
36	107.6	2.1	1268	9	AB093252	AB093252 Mus muscu
37	104.4	2.0	5191	9	AB093252	AB093252 Sequence
38	102.8	2.0	3098	6	CO715342	CO715342 Homo sapi
39	98.4	1.9	6951	8	AB014522	AB014522 Homo sapi
40	94.2	1.8	1189	8	HS288057	HS288057 Homo sapi
41	92	1.8	608	6	CQ527362	CQ527362 Sequence
42	90	1.7	627	5	BX933073	BX933073 Gallus ga
43	89.8	1.7	5770	8	HSM809084	HSM809084 Homo sapi
44	86.6	1.7	5905	8	AK074338	AK074338 Homo sapi
45	81.8	1.6	3141	2	AK116783	AK116783 Clona int

ORIGIN

Query Match	100.0%	Score 5145	DB 6	Length 5145
Best local similarity	100.0%	Pred. No. 0	Indels 0	Gaps 0
Matches 5145	Conservative 0	Mismatches 0		
QY	1	GAATTGGGACGAGCGCGCTCCGAAATTTGTTGGCGGAGACGAGAGCGAATGACATTG	60	
DB	1	GAATTGGGACGAGCGCGCTCCGAAATTTGTTGGCGGAGACGAGAGCGAATGACATTG	60	
QY	61	CAACGATGGCTATCGGAGCCGACGACCTGTGATGCTTTATCCAGCAANTGCCAAGG	120	
DB	61	CAACGATGGCTATCGGAGCCGACGACCTGTGATGCTTTATCCAGCAANTGCCAAGG	120	
QY	121	CGGACATGCTGTGAAGGTACAGTGTGCGGAGATCGTGAATCTCTTAGCCACGACA	180	
DB	121	CGGACATGCTGTGAAGGTACAGTGTGCGGAGATCGTGAATCTCTTAGCCACGACA	180	
QY	181	CAAACTGAATGTGTGACGACGACATGAGATTCTTATGACGTTGATGCGATGCTGA	240	

Db 181 CAAATCTAATGTGTGCAAGGACATGGGATTCCTTAATGACCGTTGATGATCATGCTGGA 240
Qy 241 CGGCGAGCCACTTTAAGATTGACAAAAGTCCCTGAGGCGTTCTCGAGCTAATAAAGC 300
Db 241 CGGCGAGCCACTTTAAGATTGACAAAAGTCCCTGAGGCGTTCTCGAGCTAATAAAGC 300
Qy 301 GATTGGGAGCGGATTTTAATGCAATACCGGCTACCGTTCTGCGACATGATGATCGAGC 360
Db 301 GATTGGGAGCGGATTTTAATGCAATACCGGCTACCGTTCTGCGACATGATGATCGAGC 360
Qy 361 TGGGAGACAGCGGAGACACAGTCCGCGAGAAAGCGCACTTCTGCTGGCGACCTCAAGC 420
Db 361 TGGGAGACAGCGGAGACACAGTCCGCGAGAAAGCGCACTTCTGCTGGCGACCTCAAGC 420
Qy 421 AGCAGAGAGTGTCTCGCGCCGAGGCGCTGATGACAAAGTGGGCACTAGCTGCTTCAAGC 480
Db 421 AGCAGAGAGTGTCTCGCGCCGAGGCGCTGATGACAAAGTGGGCACTAGCTGCTTCAAGC 480
Qy 481 ACAAGAACGCGCAAGGTGCGCGAGAGATTCTTCAAGACATTTGTGAACGCTCTTCATGAGT 540
Db 481 ACAAGAACGCGCAAGGTGCGCGAGAGATTCTTCAAGACATTTGTGAACGCTCTTCATGAGT 540
Qy 541 ACGGACCCAGCAGCTTGTGTGCTGCTGTATATACACAGTTGTGCACTTTCGAGC 600
Db 541 ACGGACCCAGCAGCTTGTGTGCTGCTGTATATACACAGTTGTGCACTTTCGAGC 600
Qy 601 ATCCACAGTTAATGATGAGGAGCGGCGCATCCAAAGCTAGTGGAAATCTAACAGCATG 660
Db 601 ATCCACAGTTAATGATGAGGAGCGGCGCATCCAAAGCTAGTGGAAATCTAACAGCATG 660
Qy 661 TAGGGGATCGATTGCGCCGAGACCTCCGTCGATGAGCATGTTCTGCTGAAATTTGG 720
Db 661 TAGGGGATCGATTGCGCCGAGACCTCCGTCGATGAGCATGTTCTGCTGAAATTTGG 720
Qy 721 CTATGTTGGACCAAAAGTTGCCAGCGATGCAAAAGAGGCTACTGCTACTTTCAGGCC 780
Db 721 CTATGTTGGACCAAAAGTTGCCAGCGATGCAAAAGAGGCTACTGCTACTTTCAGGCC 780
Qy 781 TTAATAAACACGAATGCAATGAGTGGGCTTGGACGAGGCGACAAATTGGGTTGAGGG 840
Db 781 TTAATAAACACGAATGCAATGAGTGGGCTTGGACGAGGCGACAAATTGGGTTGAGGG 840
Qy 841 AGCGACCCAGCAGATGATTAAGCGGCACTACCTGCGGCTTCTGATCACTGCGCC 900
Db 841 AGCGACCCAGCAGATGATTAAGCGGCACTACCTGCGGCTTCTGATCACTGCGCC 900
Qy 901 CAAAACCCAAATGTAACGATGTAACGATGTAACGATGTAACGATGTAACGATGTAACG 960
Db 901 CAAAACCCAAATGTAACGATGTAACGATGTAACGATGTAACGATGTAACGATGTAACG 960
Qy 961 AATCTAGCTTGAAGTGTGCTCCGCAATGGAATGGAATGGAATGGAATGGAATGGAATG 1020
Db 961 AATCTAGCTTGAAGTGTGCTCCGCAATGGAATGGAATGGAATGGAATGGAATGGAATG 1020
Qy 1021 TCTCAAGCAAGTACTAGTATCATGATTAATAAAACGCAAGTGGGAAACGTTGG 1080
Db 1021 TCTCAAGCAAGTACTAGTATCATGATTAATAAAACGCAAGTGGGAAACGTTGG 1080
Qy 1081 ATGCTCTCAAGAAATGATGAGGCTATGCTATCTACAGTATCACTAGCGCGAGTTTG 1140
Db 1081 ATGCTCTCAAGAAATGATGAGGCTATGCTATCTACAGTATCACTAGCGCGAGTTTG 1140
Qy 1141 TCGCTGTACAGTAAAGAAATTTGCTTAAAGCTTGTGGAATCTCAAGAGAACTAC 1200
Db 1141 TCGCTGTACAGTAAAGAAATTTGCTTAAAGCTTGTGGAATCTCAAGAGAACTAC 1200
Qy 1201 GATCAGAGGTATCGCGAGGCGTGTGATCACATGCTTACATGCTTAAAGCGCTGAGAA 1260
Db 1201 GATCAGAGGTATCGCGAGGCGTGTGATCACATGCTTACATGCTTAAAGCGCTGAGAA 1260
Qy 1261 ATAAACTAGATGCTTCTGCTGAGCAATTTGGAGCACTGATTAATTAATACAGAA 1320
Db 1261 ATAAACTAGATGCTTCTGCTGAGCAATTTGGAGCACTGATTAATTAATACAGAA 1320

Db 1321 ATAAACTAGATGCTTCTGCTGAGCAATTTGGAGCACTGATTAATTAATACAGAA 1320
Qy 1321 GCGGAAAGGTCAATGATCCGCTTCCAAATAGCTGGAATATATCATTAAGTATAC 1380
Db 1321 GCGGAAAGGTCAATGATCCGCTTCCAAATAGCTGGAATATATCATTAAGTATAC 1380
Qy 1381 ATGACCAAAAGTCTCTCAAGTATCAACAGACACTTGTGAATCAATCAAAAGTCAAGCA 1440
Db 1381 ATGACCAAAAGTCTCTCAAGTATCAACAGACACTTGTGAATCAATCAAAAGTCAAGCA 1440
Qy 1441 TAAGTCCACATGCTGATGAGTGTGATGAGTGTGCTTCAAGAGTGGGAGCGAGGCGT 1500
Db 1441 TAAGTCCACATGCTGATGAGTGTGATGAGTGTGCTTCAAGAGTGGGAGCGAGGCGT 1500
Qy 1501 TGAAGAGAAATGCCACCGTACTAAGGAGCACTTAAAAAATCATTGGCGATGAGACT 1560
Db 1501 TGAAGAGAAATGCCACCGTACTAAGGAGCACTTAAAAAATCATTGGCGATGAGACT 1560
Qy 1561 GCGATGCAAGCGCGCATTCAGATAGCGCTATTTGGGCTTTCAGGCGTCACTTTCAGAGC 1620
Db 1561 GCGATGCAAGCGCGCATTCAGATAGCGCTATTTGGGCTTTCAGGCGTCACTTTCAGAGC 1620
Qy 1621 TGGGAGATCAAAATATGGAACATTAAGCAATAGTGGCCAGCGGCACTTGAAGAGGAGC 1680
Db 1621 TGGGAGATCAAAATATGGAACATTAAGCAATAGTGGCCAGCGGCACTTGAAGAGGAGC 1680
Qy 1681 GAGAGGCGCGGAGAGAGAGAACTGTACTGAGGACTGAGGACTGAGGACTGAGGACTGAG 1740
Db 1681 GAGAGGCGCGGAGAGAGAGAACTGTACTGAGGACTGAGGACTGAGGACTGAGGACTGAG 1740
Qy 1741 GCATGATATCCCGCATTTGGCGGAAACCTGGAACCTTGCAGAAAGCCGCTTGTATGA 1800
Db 1741 GCATGATATCCCGCATTTGGCGGAAACCTGGAACCTTGCAGAAAGCCGCTTGTATGA 1800
Qy 1801 GATCATTTTCAAGGCGGTGGAACATGCGGCTGCTCAAGAGCGCAAGTTAGAGCGCAATTA 1860
Db 1801 GATCATTTTCAAGGCGGTGGAACATGCGGCTGCTCAAGAGCGCAAGTTAGAGCGCAATTA 1860
Qy 1861 CACTATATTCGAGCAAAAGAAACCTTTAGAACCTTAATTAATTCGAACGAGCATGATGA 1920
Db 1861 CACTATATTCGAGCAAAAGAAACCTTTAGAACCTTAATTAATTCGAACGAGCATGATGA 1920
Qy 1921 CAGAGACAGCAGCGATGATCACTGCGCAGACCTCGCTGGAATTTCCATAGCGGTGCA 1980
Db 1921 CAGAGACAGCAGCGATGATCACTGCGCAGACCTCGCTGGAATTTCCATAGCGGTGCA 1980
Qy 1981 CACCAAGTAAACGCGGAGATCGGTTACACACGCGCCGCGGAGCGGAGAGATATCTC 2040
Db 1981 CACCAAGTAAACGCGGAGATCGGTTACACACGCGCCGCGGAGCGGAGAGATATCTC 2040
Qy 2041 AGTCAAGCAGAGATCGAGATCCACTCGCCAAAGCAAAATCTGAGGATCAATACGTTG 2100
Db 2041 AGTCAAGCAGAGATCGAGATCCACTCGCCAAAGCAAAATCTGAGGATCAATACGTTG 2100
Qy 2101 GATTTGTAATTAATCAATGCTGGGCGCACTGCGCGCATACCCAAAAAAGCCTCCGGAATAC 2160
Db 2101 GATTTGTAATTAATCAATGCTGGGCGCACTGCGCGCATACCCAAAAAAGCCTCCGGAATAC 2160
Qy 2161 CCGGAGCAGAGCAGCTAGGGAACGAGTCCAACTAGGTCAGGTTGAGTGTGATGA 2220
Db 2161 CCGGAGCAGAGCAGCTAGGGAACGAGTCCAACTAGGTCAGGTTGAGTGTGATGA 2220
Qy 2221 AACGAGTATGATCTTACAGGTGCGGGGCTTCGACGTACGCGCCGAGAGAAACACCCAG 2280
Db 2221 AACGAGTATGATCTTACAGGTGCGGGGCTTCGACGTACGCGCCGAGAGAAACACCCAG 2280
Qy 2281 TAAAGCAATGCGGCGCGGCGCACTGTGCGCAATCCGTAAGAGCAATATCATTTAG 2340
Db 2281 TAAAGCAATGCGGCGCGGCGCACTGTGCGCAATCCGTAAGAGCAATATCATTTAG 2340
Qy 2341 GCGTCGAGATGATGAGACAGACGATATTTCCGCGGAGCTACATGCGAGCGCGGAA 2400
Db 2341 GCGTCGAGATGATGAGACAGACGATATTTCCGCGGAGCTACATGCGAGCGCGGAA 2400

Db	4561	TAGATTCCTGCTCGTGCCACCACAAAACAAGACACAGACGCGGTTCCTCCCTCGGCTG	4620
Qy	4621	AGAAAGAAAGTGAAGAGCGCGCGGACATTAAATATATATTTATTAACATTAACTTAATTTTA	4680
Db	4621	AGAAAGAAAGTGAAGAGCGCGCGGACATTAAATATATATATTTATTAACATTAAATTTA	4680
Qy	4681	TAACTAATTTATTAACCGATCATCCGTTACTGTGTGTAACGTCGACGATATATTTACAGAG	4740
Db	4681	TAACTAATTTATTAACCGATCATCCGTTACTGTGTGTAACGTCGACGATATATTTACAGAG	4740
Qy	4741	ATGCAATATGCGGCCCGGAAAGAAAGTCAAAGGCGCATGGCCATCTATATGTAGACGAAAAA	4800
Db	4741	ATGCAAAATGCGGCCCGGAAAGAAAGTCAAAGGCGCATGCCCATCTATATGTAGACGAAAAA	4800
Qy	4801	CTATTTTATACATPAAACGGGAACCTAAAGGAGTAAATCCGCAAAAGTATCTAAATTTGAG	4860
Db	4801	CTATTTTATACATPAAACGGGAACCTAAAGGAGTAAATCCGCAAAAGTATCTAAATTTGAG	4860
Qy	4861	CCAACCTCGCATCCTCACTTCTGCATCAGTGTGACCTCTTTTATTCAGAGCTCCTAACTA	4920
Db	4861	CCAACCTCGCATCCTCACTTCTGCATCAGTGTGACCTCTTTTATTCAGAGCTCCTAACTA	4920
Qy	4921	TTAAACGCTGATCTGTTTAAATTTGTGAAGCCTATATACCGCTCTTATATGGAACCTCAGCG	4980
Db	4921	TTAAACGCTGATCTGTTTAAATTTGTGAAGCCTATATACCGCTCTTATATGGAACCTCAGCG	4980
Qy	4981	TGAGCTTTTACACCTTGTTCACGGCAGCTTGATACCTTATTTTCGTTCCGCTCAACCTA	5040
Db	4981	TGAGCTTTTACACCTTGTTCACGGCAGCTTGATACCTTATTTTCGTTCCGCTCAACCTA	5040
Qy	5041	AATTTCTAATTACTATTTATTAATAGATTTCTGCTCTTACACCTGAACATTTTGTAAATTA	5100
Db	5041	AATTTCTAATTACTATTTATTAATAGATTTCTGCTCTTACACCTGAACATTTTGTAAATTA	5100
Qy	5101	ATTTTAATTAACAAGCCACACCAAGATTTAAATTCATATAAAAA	5145
Db	5101	ATTTTAATTAACAAGCCACACCAAGATTTAAATTCATATAAAAA	5145

```

RESULT 2
LOCUS AB031048
DEFINITION AB031048 5959 bp mRNA linear INV 07-APR-2000
Drosophila melanogaster orbit mRNA for microtubule
associated-protein orbit, complete cds.
ACCESSION AB031048
VERSION AB031048.1 GI:7527325
KEYWORDS microtubule associated-protein orbit.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (sites)
AUTHORS Inoue,Y.H., do Carmo Avides,M., Shiraki,M., Deak,P., Yamaguchi,M.,
Nishimoto,Y., Matsukage,A. and Glover,D.M.
TITLE Orbit, a novel microtubule-associated protein essential for mitosis
in Drosophila melanogaster
J. Cell Biol. 149 (1), 153-166 (2000)
10747094
2 (bases 1 to 5959)
Inoue,Y.
Direct Submission
Submitted (11-AUG-1999) Yoshihiro H Inoue, Kyoto Institute of
Technology, Drosophila Genetic Resource Centre, Matsugasaki, Sakyo,
Kyoto 606-8585, Japan (E-mail:yhinoue@rochoan.bio.kit.ac.jp,
tel:+81(75)724-7788, fax:+81(75)724-7710)
FEATURES
     source          Location/Qualifiers
1..5959
     /organism="Drosophila melanogaster"
     /mol_type="mRNA"
     /db_xref="taxon:7227"
     /chromosome="3"

```

gene
CDS

polya_site

ORIGIN

```
1. 5959
/gene="orb1t"
770. 5248
/gene="orb1t"
/codon_start=1
/product="microtubule associated-protein orb1t"
/db_xref="GI:7527326"
/translation="MAIRKPSDDGFIQOMPKADMVKVQDAEDVTFVLSDDTNSITVC
TDWFGILIDGLMPWLTGSHFKIAGKSLSEAFSELIKRGSDNATATATVLPVHIDRLGSG
RDVREKQGLMLDLMDLMEHFKVLPQALIDKATSCFKPKAKVEEFLQTVMLAHEG
TOOLSVAVYIPVPCALIGDPTVNVREALIQTVLEIYHVGDRIPDIRMDMDVPAVSL
AMLEOKEDVOKGGLLPSALKNTNGNGVGLDADNIGLDERPTMRKRLHSAVSS
LRKPNVNDVADGAVTMESPESSEFVQQLINFAKMDWDIYKQVLIISPDNAWV
EKRVDAIKTKIRALLIISYHPOPOFAVOLKEISLSPDILKEELRSQVIREACTIYV
MSTTLNRKLDAPCWSLIEHLINLIQNSAKVIASASTALKYITIKYTHAPLKYITDTT
LNSOKSRSTLCELMVLLFEEWQTALEKRNATVLDTLKSIQGDADCAHRSRIA
YMFRRHFPELDQIIGTDLIAQRALERRERBGGGGTGTGTAPETRTVRSIOR
TPEQTLKPTSPMSKSIADVTAARAKVRAQYVLYRQRPDLPNNMSQAMTGGAAN
GSLPRRLNNSGGPTPATPGVTPPRPGAGVGSOSGGRSSTPSKLDGOCGLIEN
VYNGATGALPKKASGILPRSTASSRETSIPRTASVYSGAGSRPDENNPDR
PAPPAULLAOSREAEHTLVGDDGQEPYVSGDYMREGGRMKMKMLRDSDDIPSA
SVSCSERSPDSSTYRKNKSVSLSGSHTRLDWSTQAPPDITITIOFCASHMSEK
DGIILSTQYLADEKELTQOQKCVLDMFRFMENDHTKYVSLQDLYTELLIVHANET
SRNGSSSCFLRLFNKGLGTDLNLNMSKIKWTLQVNHVEFPTQLKELFEIISDTOT
PTTQTALIRFLDLYANTYCKSDSPDSQACERTVLGLAOLAODKMEIORSAR
SCVALYINNTPOMTILLADI.PVYODSARSCISHMRROSQNSGANSRSPSSS
SSKRPLOSRSEVCPFASLQSHHOLDISSTSPRROSQVBEELPSELDIHNIOKRS
BEIRHCFGGQYOTALANPENGHQLQTHDQSGQSLNSKTSQASANTQSGTTPES
ATMRDLNERERTTQNAKSPGDDAKAVTVISINAEENGELIASNLMESEVRAVLTUT
KQDPVELLQTSLTNLGICIKGNCELPNKHFRSIMBLNITLEAHTDVVIAQIHLVS
KIKRSNMRHNMHHFLELILKKIIOCCYQSKELARDIDSMIPRIASLPDLSINIVN
PVATSEFPTNLCAIKILLEVTBHGSEIIDLALDVFPEMLASADTQOMVKAFAV
CIYKLFVJGEEKVEVRKSLTNBPKYRLNVTYIEKQRCNCSGGGSSSTKNSAASSS"
5916
/gene="orb1t"
```

Query Match	Best Local Similarity	Score	DB 2	Length
99.5%	99.8%	5120.8	5959	
Matches 5128	Conservative 0	Mismatches 12	Indels 0	Gaps 0

Query	6	710	66	770	126	1830	186	890	246	950	306	1010	366	1070	426
DGCGACGAGCGCGCTCCGAAATTTGTTGGCCGAGACGAGAGCGAATAGCATTTGCACG	CGCGACGAGCGCGCTCCGAAATTTGTTGGCCGAGACGAGAGCGAATAGCATTTGCACG	CACGCGACGCGCGGATTTGCAAAATATGTTGCGCGACAGAGAGGAGCAATAGCATTTGCACG	ATGGCCATATCGAAGGCCGACGCGACCTGAGATGGCTTTATCCAGCAAAATGCCAAGGGCGAC	ATGACCATATCGAAGGCCGACGCGACCTGAGATGGCTTTATCCAGCAAAATGCCAAGGGCGAC	ATGCGTGTGAAGGTATACAGCTCGCGAGAGATCTGTGTACATTTCTTAGCGACACACAAAC	ATGCGTGTGAAGGTATACAGCTCGCGAGAGATCTGTGTACATTTCTTAGCGACACACAAAC	TCAATTTGCTGCACGCGACATGGGATTCCTTATTTAGCGGTTTATGTCATGGCTGACGGGC	TCAATTTGCTGCACGCGACATGGGATTCCTTATTTAGCGGTTTATGTCATGGCTGACGGGC	AGCCACTTTTAAGTTTGCACAAAAGTCCCTGAGAGCGTTCTCGGAGCTAATAAGCGATTG	AGCCACTTTTAAGTTTGCACAAAAGTCCCTGAGAGCGTTCTCGGAGCTAATAAGCGATTG	GGCAGCGATTTTAATGCATACAGCGCTACCGTTCTGCGACATGTGATCGATCGGCTGGGA	GGCAGCGATTTTAATGCATACAGCGCTACCGTTCTGCGACATGTGATCGATCGGCTGGGA	GACAGCAGGAGACACAGTCCGCGAGAGAGGCGCACTTCTGTGCGACACCTCATGAGACAC	GACAGCAGGAGACACAGTCCGCGAGAGAGGCGCACTTCTGTGCGACACCTCATGAGACAC	AGAGTGCCTTCGCGCCGACGGCGCTGATCGACAAAGCTGGCCATTAAGCTGCTTCAAGCACAG
485	485	769	125	829	185	889	245	949	305	1009	365	1069	425	1129	485

Db 1130 AGAGTCTTCCGCCAGCGCTGATGACAGCTGGCCACTAGCTTCTTCAAGACAAG 1189
QY 486 AACGCCAAGGTGGGAGGAGGATTCCTTACAGCATTTGTGAACGCTCTTCACTAGTACGG 545
Db 1190 AAGGCCAAGGTGGGAGGAGGATTCCTTACAGCATTTGTGAACGCTCTTCACTAGTACGG 1249
QY 546 ACCCAGCAGTTAGTGTTCGCTCTATATACACAGTTGTGACACTTCTCGAGATCCC 605
Db 1250 ACCCAGCAGTTAGTGTTCGCTCTATATACACAGTTGTGACACTTCTCGAGATCCC 1309
QY 606 ACAATTATGTGAGGAGGAGCGCCATCCAAACCTTATGTGAATCTAACAGCATGTAGG 665
Db 1310 ACAGTTAATGTGAGGAGGAGCGCCATCCAAACCTTATGTGAATCTAACAGCATGTAGG 1369
QY 666 GATCGATTGGGCCCAACCTCCGTCGATGAGATGTTCTCCGCTCGAAATTGGGCTATG 725
Db 1370 GATCGATTGGGCCCAACCTCCGTCGATGAGATGTTCTCCGCTCGAAATTGGGCTATG 1429
QY 726 TTGGAGCAAAAGTTCCACACAGTCAAAAGAGGGTCTACTGCTTACCTTCAGCCTTAA 785
Db 1430 TTGGAGCAAAAGTTCCACACAGTCAAAAGAGGGTCTACTGCTTACCTTCAGCCTTAA 1489
QY 786 AACACGAATGGAATGAGATGAGGCTTGAAGAGCGGACCAATATTTGGTTGAGGAGCG 845
Db 1490 AACACGAATGGAATGAGATGAGGCTTGAAGAGCGGACCAATATTTGGTTGAGGAGCG 1549
QY 846 CCCACACAGATGATTTAAGCGGCCACTACACTCGGCGCTTTCGTCATCACTGCGCCAAA 905
Db 1550 CCCACACAGATGATTTAAGCGGCCACTACACTCGGCGCTTTCGTCATCACTGCGCCAAA 1609
QY 906 CCCAATGTGACGATGTGACCGGTGATGCGGCGCGCTTAAACATGTGAATCTTTCGATCT 965
Db 1610 CCCAATGTGACGATGTGACCGGTGATGCGGCGCGCTTAAACATGTGAATCTTTCGATCT 1669
QY 966 AGCTTTGAGTGTCCCGCAATTGAAATTTTCAAGCTTAAGACATGACATATCTAC 1025
Db 1670 AGCTTTGAGTGTCCCGCAATTGAAATTTTCAAGCTTAAGACATGACATATCTAC 1729
QY 1026 AAGCAATCTATGATGATCACTGATTAATAAGCAGACTGGAGAAAGCTGTGATGCT 1085
Db 1730 AAGCAATCTATGATGATCACTGATTAATAAGCAGACTGGAGAAAGCTGTGATGCT 1789
QY 1086 CTCGAAGAATCAGGAGCATTTGCTCATTTCTCAGCTATCACTCAGCGCGAGTTTGTGCT 1145
Db 1790 CTCGAAGAATCAGGAGCATTTGCTCATTTCTCAGCTATCACTCAGCGCGAGTTTGTGCT 1849
QY 1146 GTACAGCTAAAGAAATTTGCTTAAGCTTGTGAGCACTCTCAAGAGAGAACTACATCA 1205
Db 1850 GTACAGCTAAAGAAATTTGCTTAAGCTTGTGAGCACTCTCAAGAGAGAACTACATCA 1909
QY 1206 CAGGTATCCGAGGCGGTGATCACTATGCTTCACTGATGAGCGCTGAAGAAATTA 1265
Db 1910 CAGGTATCCGAGGCGGTGATCACTATGCTTCACTGATGAGCGCTGAAGAAATTA 1969
QY 1266 CTGATGAGCTTCTGCTGAGACATTTTGAAGACCTCTGATTAATTAATACAGACGCG 1325
Db 1970 CTGATGAGCTTCTGCTGAGACATTTTGAAGACCTCTGATTAATTAATACAGACGCG 2029
QY 1326 AAGGTATGTCATCCGCTTCCACAAATAGCTCTGAAGTATCATTAAGTATACATGCA 1385
Db 2030 AAGGTATGTCATCCGCTTCCACAAATAGCTCTGAAGTATCATTAAGTATACATGCA 2089
QY 1386 CCAAAGCTGTCGAAGATCTACAGACACTGTGAATCAATCAAGTGAAGGACATAGG 1445
Db 2090 CCAAAGCTGTCGAAGATCTACAGACACTGTGAATCAATCAAGTGAAGGACATAGG 2149
QY 1446 TCCACACTGTGATGATGATGCTCTTTCGAGAGAGGAGAGAGAGGCTTGA 1505
Db 2150 TCCACACTGTGATGATGATGCTCTTTCGAGAGAGGAGAGAGAGGCTTGA 2209
QY 1506 AGGAATGCCACCGTACTAAGGAGACCTTAAATAATTCATTGGCGATGCGAT 1565
Db 2210 AGGAATGCCACCGTACTAAGGAGACCTTAAATAATTCATTGGCGATGCGAT 2269

QY 1566 GCACGCCCATTCAGATACGCTATTGGCTTTGAGCGTCACTTTCAGAGCTGGC 1625
Db 2270 GCACGCCCATTCAGATACGCTATTGGCTTTGAGCGTCACTTTCAGAGCTGGC 2329
QY 1626 GATCAATTTATGGAACATTTAATCATGATGCTCCAGCGGCACTTGAAGGGAACAGAG 1685
Db 2330 GATCAATTTATGGAACATTTAATCATGATGCTCCAGCGGCACTTGAAGGGAACAGAG 2389
QY 1686 GCGCGCGAGGAGGAGAACTGGTACTGGGACTGGGACTGCACTGAAACGAGACGCACT 1745
Db 2390 GCGCGCGAGGAGGAGAACTGGTACTGGGACTGGGACTGCACTGAAACGAGACGCACT 2449
QY 1746 GATCCCGCATTTGGCGGAACACTTGAACCTTGAAGCCCAACGCTTATGATGATCC 1805
Db 2450 GATCCCGCATTTGGCGGAACACTTGAACCTTGAAGCCCAACGCTTATGATGATCC 2509
QY 1806 ATTTCAAGCGGTGACACTGCGGCTGCTCAACGAGCCAAAGTTAGAGCGCAATTAACCTA 1865
Db 2510 ATTTCAAGCGGTGACACTGCGGCTGCTCAACGAGCCAAAGTTAGAGCGCAATTAACCTA 2569
QY 1866 TATTCAGGCAAAAGGAACCTTTAGAACCTTAATTAATTCACAGGCACTGATGACAGA 1925
Db 2570 TATTCAGGCAAAAGGAACCTTTAGAACCTTAATTAATTCACAGGCACTGATGACAGA 2629
QY 1926 GCAGCAGCAGTGAATCACTGCGCAGACCTCGCTGAATTCGAATAGCGGTGGCACCA 1985
Db 2630 GCAGCAGCAGTGAATCACTGCGCAGACCTCGCTGAATTCGAATAGCGGTGGCACCA 2689
QY 1986 GCTTACAACCGCGGATTCGTTACACCAACGCGCCGCGGAGCGGACAGATATCTCAGTCA 2045
Db 2690 GCTTACAACCGCGGATTCGTTACACCAACGCGCCGCGGAGCGGACAGATATCTCAGTCA 2749
QY 2046 CAGCAGGATGAGATCCACCTCGCCAGACAAACCTGAGGATCATGACGGTGTATT 2105
Db 2750 CAGCAGGATGAGATCCACCTCGCCAGACAAACCTGAGGATCATGACGGTGTATT 2809
QY 2106 GGTAAATTACTATGTTGGGCACTGGCGCATACCCAAAGAGCTTCGGAAATACCCCA 2165
Db 2810 GGTAAATTACTATGTTGGGCACTGGCGCATACCCAAAGAGCTTCGGAAATACCCCA 2869
QY 2166 AGACACAGCAGCTCTAGGGAACGAGTCCAACTAGGTGAGTGTGCTTGAATGAACGC 2225
Db 2870 AGACACAGCAGCTCTAGGGAACGAGTCCAACTAGGTGAGTGTGCTTGAATGAACGC 2929
QY 2226 AGTATCTACTCTACAGGTGCGGGTCTGACGTAAGCGCGAGAGAAACACAGTAA 2285
Db 2930 AGTATCTACTCTACAGGTGCGGGTCTGACGTAAGCGCGAGAGAAACACAGTAA 2989
QY 2286 CCATCGCGCGCGGCAAGCTGCTGGGCAATCCCGTGAAGCAGAACATACATTAGGCGTC 2345
Db 2990 CCATCGCGCGCGGCAAGCTGCTGGGCAATCCCGTGAAGCAGAACATACATTAGGCGTC 3049
QY 2346 GGAATGATGAGCAGCAGCACTATGTTTCCGGGACTAATGCGAGCGCGGAATGCGG 2405
Db 3050 GGAATGATGAGCAGCAGCACTATGTTTCCGGGACTAATGCGAGCGCGGAATGCGG 3109
QY 2406 ATGGTGAAGAACTCATGGAAGGTGATGATGATGATGATGATGATGATGATGATGATGAT 2465
Db 3110 ATGGTGAAGAACTCATGGAAGGTGATGATGATGATGATGATGATGATGATGATGATGAT 3169
QY 2466 GTGTGTTCAAGACGATCTTTCAGCTCAGCTACATGAGGATTAATTCGAATCTCA 2525
Db 3170 GTGTGTTCAAGACGATCTTTCAGCTCAGCTACATGAGGATTAATTCGAATCTCA 3229
QY 2526 CTTAGCGGAGGCAACCCGCTTGAATGAGAGACGACGCGCGGCACTTTGACGACATC 2585
Db 3230 CTTAGCGGAGGCAACCCGCTTGAATGAGAGACGACGCGCGGCACTTTGACGACATC 3289
QY 2586 GAGAGATTAATTCAGTTGCGCATGACGACATTTGGTGAAGGAAGATGCGCTGATC 2645
Db 3290 GAGAGATTAATTCAGTTGCGCATGACGACATTTGGTGAAGGAAGATGCGCTGATC 3349

QY 2646 AGCCTCACAGATATCTGGCCGATGAGAAAGAGCTCACTCAGACGCACTTAAATGCGTT 2705
 Db 3350 AGCCTCACAGATATCTGGCCGATGAGAAAGAGCTCACTCAGACGCACTTAAATGCGTT 3409
 QY 2706 CTGACATGTTCCGAGAGATGTTCAATGATACACACACCAAGGTACTCTGTTTTTG 2765
 Db 3410 CTGACATGTTCCGAGAGATGTTCAATGATACACACACCAAGGTACTCTGTTTTTG 3469
 QY 2766 GACACGGTCACCGAATGATCTGGTTGATGCGAAGCACTTACGAAATGGCTTTCA 2825
 Db 3470 GACACGGTCACCGAATGATCTGGTTGATGCGAAGCACTTACGAAATGGCTTTCA 3529
 QY 2826 TCTGTTTGAAGCGCGGTGTTCAACAAATTTGGGCAACGATCTACTCAATTCGATGACAGC 2885
 Db 3530 TCTGTTTGAAGCGCGGTGTTCAACAAATTTGGGCAACGATCTACTCAATTCGATGACAGC 3589
 QY 2886 AAGATTGGAAGAGCTGAGAGTGTCCACGAATATTTCCCGACGAGTTGCACTTTAA 2945
 Db 3590 AAGATTGGAAGAGCTGAGAGTGTCCACGAATATTTCCCGACGAGTTGCACTTTAA 3649
 QY 2946 GAACTATTAGATCATTCGGATTCTACTCAACGCGCACTACCAAGACGCGATCGCC 3005
 Db 3650 GAACTATTAGATCATTCGGATTCTACTCAACGCGCACTACCAAGACGCGATCGCC 3709
 QY 3006 ATTCTGCGCTTCTTACGGATCTGGCCAAATACATCTGTAAAGCAGGACTTTCCAGC 3065
 Db 3710 ATTCTGCGCTTCTTACGGATCTGGCCAAATACATCTGTAAAGCAGGACTTTCCAGC 3769
 QY 3066 GACCAAGGCCAGGCTCGGAGAGCGATCCTAAAGCTGGCCGACGTCGGCGGATCAG 3125
 Db 3770 GACCAAGGCCAGGCTCGGAGAGCGATCCTAAAGCTGGCCGACGTCGGCGGATCAG 3829
 QY 3126 AAGTCAGTGAAGTGCCTCCAGGCGAGAGCTGAGTGGCCCTGTATACCTGAAT 3185
 Db 3830 AAGTCAGTGAAGTGCCTCCAGGCGAGAGCTGAGTGGCCCTGTATACCTGAAT 3889
 QY 3186 ACCCGCAATGACCTTTTACTGCGCGACCTGCAAGGTATATCAGAGACTGCGCGA 3245
 Db 3890 ACCCGCAATGACCTTTTACTGCGCGACCTGCAAGGTATATCAGAGACTGCGCGA 3949
 QY 3246 TCCTCGATCCATTTGGACATGAGGCGGCAAGCCAAAGTTGCAATTTGGGCGCAATTCG 3305
 Db 3950 TCCTCGATCCATTTGGACATGAGGCGGCAAGCCAAAGTTGCAATTTGGGCGCAATTCG 4009
 QY 3306 CTTAGTAGCTCTCCATTGAGCAGTAGCAGTCCCAAGCTTTGCAAGTCCCTGTGGGT 3365
 Db 4010 CTTAGTAGCTCTCCATTGAGCAGTAGCAGTCCCAAGCTTTGCAAGTCCCTGTGGGT 4069
 QY 3366 CCATTTGCTTGTTCAGAGCCACCAACCACTTAGCATCAGCTTACTTATGTCACGC 3425
 Db 4070 CCATTTGCTTGTTCAGAGCCACCAACCACTTAGCATCAGCTTACTTATGTCACGC 4129
 QY 3426 TCCCGGAGTCTTCCGTTGAGCAGAGAGTGTCTTTTCTCGGAGCTGAGATTCACAGC 3485
 Db 4130 TCCCGGAGTCTTCCGTTGAGCAGAGAGTGTCTTTTCTCGGAGCTGAGATTCACAGC 4189
 QY 3486 AACATTGAGAAAGTGGGAGAGATCCGAGCAGTCTTGGCGGTACAGTACAGAGCG 3545
 Db 4190 AACATTGAGAAAGTGGGAGAGATCCGAGCAGTCTTGGCGGTACAGTACAGAGCG 4249
 QY 3546 CTGGGCGCCAAATGCTTCAATGACACTTGGAGTATCCATGACAGGCGCAACAGATTCG 3605
 Db 4250 CTGGGCGCCAAATGCTTCAATGACACTTGGAGTATCCATGACAGGCGCAACAGATTCG 4309
 QY 3606 TGTGATATCCCTGTTTCAATCCAGAGCGCAATCTGCGGCAACATCCCACTCAAT 3665
 Db 4310 TGTGATATCCCTGTTTCAATCCAGAGCGCAATCTGCGGCAACATCCCACTCAAT 4369
 QY 3666 ACACCTGATGAGCAACATGAGCTGATATCTGAGCGGAGAAAGACCACTCAGAAC 3725
 Db 4370 ACACCTGATGAGCAACATGAGCTGATATCTGAGCGGAGAAAGACCACTCAGAAC 4429
 QY 3726 GCCAAGTCACCACTGACGATGCCAAGGTGATCAAGGTCTGATTAATATGCGTGAAT 3785

Db 4430 GCCAAGTCACCACTGACGATGCGCAAGGTGATCAGGTCCGATTAATATGCGTGAAT 4489
 QY 3786 GAGAGCTGATATCTGGCCAGCAACCTGATGAGAGGAAAGTGTGCGTGTGACG 3845
 Db 4490 GAGAGCTGATATCTGGCCAGCAACCTGATGAGAGGAAAGTGTGCGTGTGACG 4549
 QY 3846 CTAAACAAGATCAGCCCGTGTGAGTCTTCAAGCTCACTTAATACTGGGGAATTCG 3905
 Db 4550 CTAAACAAGATCAGCCCGTGTGAGTCTTCAAGCTCACTTAATACTGGGGAATTCG 4609
 QY 3906 ATCAAGGCGGAAACTGTGAGTGTGCGCAATATGACCTTTAGATGATCAAGCAATCTG 3965
 Db 4610 ATCAAGGCGGAAACTGTGAGTGTGCGCAATATGACCTTTAGATGATCAAGCAATCTG 4669
 QY 3966 CTTAATCTTGAAGCGGAGCATACGAGCTGTATCGCTGGCTGACGTCAGT 4025
 Db 4670 CTTAATCTTGAAGCGGAGCATACGAGCTGTATCGCTGGCTGACGTCAGT 4729
 QY 4026 AAGATTATGAGAGCAACAAATGCGTCAACACTGATGCACTTTTGAAGCTGATTTG 4085
 Db 4730 AAGATTATGAGAGCAACAAATGCGTCAACACTGATGCACTTTTGAAGCTGATTTG 4789
 QY 4086 CTGAAGTATCTCAGTGTCTATCAACACAGCAAGAGGCTTTGGGGATATCGACTCGATG 4145
 Db 4790 CTGAAGTATCTCAGTGTCTATCAACACAGCAAGAGGCTTTGGGGATATCGACTCGATG 4849
 QY 4146 ATACCAAGATACACCATCTTGTCTGTGATCTGTCTCAACATTTGCAATCCGTA 4205
 Db 4850 ATACCAAGATACACCATCTTGTCTGTGATCTGTCTCAACATTTGCAATCCGTA 4909
 QY 4206 ATTGCGAGGAGTGTTCACAAATTTGTGCGCATCAAGATCTGTGAAATGAC 4265
 Db 4910 ATTGCGAGGAGTGTTCACAAATTTGTGCGCATCAAGATCTGTGAAATGAC 4969
 QY 4266 GAACACCATGTTGCGAGATCAACAGAGCCCACTTGGACATTTGTTCCCAATTCGGCG 4325
 Db 4970 GAACACCATGTTGCGAGATCAACAGAGCCCACTTGGACATTTGTTCCCAATTCGGCG 5029
 QY 4326 CGATCGGCGAGCAGACAGCAATGATGATGCGAAAGCTGCTTTCATCTGTCAG 4385
 Db 5030 CGATCGGCGAGCAGACAGCAATGATGATGCGAAAGCTGCTTTCATCTGTCAG 5089
 QY 4386 CTGTACTTGTGTTGCGGCGAGAGAGTTCAGCCGAAAGCTGTGATGCTTAATCCAGC 4445
 Db 5090 CTGTACTTGTGTTGCGGCGAGAGAGTTCAGCCGAAAGCTGTGATGCTTAATCCAGC 5149
 QY 4446 AAGTTAAGTCTCTCAACGTGTATCATCGAAGCAAGCGAACTGATGAGTGGGGAGGA 4505
 Db 5150 AAGTTAAGTCTCTCAACGTGTATCATCGAAGCAAGCGAACTGATGAGTGGGGAGGA 5209
 QY 4506 AGCTTACAAAGAACTCTTCCGCGGAGTGTCTCATGATTTGGGGAGCCCTTAATGGA 4565
 Db 5210 AGCTTACAAAGAACTCTTCCGCGGAGTGTCTCATGATTTGGGGAGCCCTTAATGGA 5269
 QY 4566 TTCTGCTGTGTCACCAACAAAGACAGAGCGGTTGTTCCCTGCGCTGAGAG 4625
 Db 5270 TTCTGCTGTGTCACCAACAAAGACAGAGCGGTTGTTCCCTGCGCTGAGAG 5329
 QY 4626 GAAGTGAAGAGCGCGGAGCACTTAATTAATTAATTAATTAATTAATTAATTAAT 4685
 Db 5330 GAAGTGAAGAGCGCGGAGCACTTAATTAATTAATTAATTAATTAATTAATTAAT 5389
 QY 4686 AATTATTAACGATCATCCGTTACTTGTGTAAGTGTGATGATATATTCAGAGATGCA 4745
 Db 5390 AATTATTAACGATCATCCGTTACTTGTGTAAGTGTGATGATATATTCAGAGATGCA 5449
 QY 4746 AATGCGGCCGAAAGAAAGTCAAGGCGCATGCGCCATCTAATGTGACAGAAATCTAAT 4805
 Db 5450 AATGCGGCCGAAAGAAAGTCAAGGCGCATGCGCCATCTAATGTGACAGAAATCTAAT 5509
 QY 4806 TATACATTAACGGGAATTAAGGAGTAAATCCGAAAGTATCTAATATTTAGCCAAC 4865

Db 5510 TATCATTAACGGGACTAAAGCGATTAATCCGCAAGTACTATAATTGACCCAC 5569

Qy 4866 TCCGCAATCTCTACTCTTCACATGAGTGTACGCTTTTTCACAGCTCTTAATTAAC 4925

Db 5570 TCCGCAATCTCTACTCTTCACATGAGTGTACGCTTTTTCACAGCTCTTAATTAAC 5569

Qy 4926 GCTATCTCTTTTAAATTGTAAGCTATATACCGCTCTTTTGTGAAGTCAAGAGTGTGC 4985

Db 5630 GCTATCTCTTTTAAATTGTAAGCTATATACCGCTCTTTTGTGAAGTCAAGAGTGTGC 5689

Qy 4986 TTTTCTACCTTGTGTTGAGGAGCTTGTATCTTATTTGCTTCCGCTCACTAAATTC 5045

Db 5690 TTTTCTACCTTGTGTTGAGGAGCTTGTATCTTATTTGCTTCCGCTCACTAAATTC 5749

Qy 5046 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5105

Db 5750 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5809

Qy 5106 AATACACAGCCCAACACAGGATTAAATCAATATTAATA 5145

Db 5810 AATACACAGCCCAACACAGGATTAAATCAATATTAATA 5849

RESULT 3

AF195498

LOCUS

DEFINITION

AF195498 5916 bp mRNA linear INV 01-NOV-2000

ras 7 (MESR7) mRNA, MESR7-3403 allele, complete cds.

ACCESSION

AF195498 GI:11066120

VERSION

AF195498.1

KEYWORDS

FLI CDNA.

SOURCE

Drosophila melanogaster (fruit fly)

ORGANISM

Drosophila melanogaster

Eukaryote; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 5916)

Huang, A.M., Rubin, G.M., Teang, G., Evans-Holm, M. and Sub, C. Full length Drosophila melanogaster CDNA sequence Unpublished

JOURNAL

2 (bases 1 to 5916)

Huang, A.M., Rubin, G.M., Teang, G., Evans-Holm, M. and Sub, C. Direct Submission

REFERENCE

Submitted (18-OCT-1999) Molecular and Cell Biology, University of California at Berkeley, 545 Life Sciences Addition Bldg., Berkeley, CA 94720-3200, USA

FEATURES

Source

1..5916

Location/Qualifiers

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/chromosome="3"

/map="78C1-2"

/clone="LD11488"

1..5916

/gene="MESR7"

/allele="3403"

771..5246

/gene="MESR7"

/allele="3403"

/note="suppressor of activated Ras when overexpressed in the adult eye"

/codon_start=1

/product="Miseexpression suppressor of ras 7"

/protein_id="AAG28470.1"

/db_xref="GI:11066121"

/translation="MAYRKPSDLDGFIQMPKADMRVKVLAEDLVTFSDPNSIVC TDWGFIDGLMPMLTGSHPFKIAQKSLAEFELIKLSDPNAYATVLPVIRLGS RDTVREKAOILRDLMEHVLPAOLIDKLSCFKKNKVEEFLOTIVNALHEG TOOLSIVYVLPVLCALIDPUNVREALIOTIVELYHVGDRRLRPDIRMDVPAKSL AMEOKSDVYKOBELLPSALKNTNGVGLDEADNIGLEBRPFRMKRPLKSHVASS LRPEPNVDYGDAGAVTMSFESSFEVQLNFKAKMDMDIYKQVLVIISDNAMW EKRVDLAKKIRALLIISYHTQPFVAVALKELSLSFVDILKEIRSGVIREACTIAY

gene

CDs

ORIGIN

Query Match 98.3%; Score 5055.6; DB 2; Length 5916;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 5108; Conservative 0; Mismatches 29; Indels 5; Gaps 3;

Qy 6 CGGACAGAGCGCGTCCGAAATTTGTCGCGCAGAACGAGCAGATGACATTTGCAACG 65

Db 711 CAGCGAGCGCGATTTGCAAAATGTTGCCGACAGAACGAGCAGATGACATTTGCAACG 770

Qy 66 ATGGCTATTTGGAAGCCGACGACCTTGATGCTTTATCCAGCAAAATGCCCAAGCGGAC 125

Db 771 ATGGCTATTTGGAAGCCGACGACCTTGATGCTTTATCCAGCAAAATGCCCAAGCGGAC 830

Qy 126 ATGCGGTGAAGGTACAGCTCGCGCAGAGATCTGATGATTTCTTACGACGACACAAC 185

Db 831 ATGCGGTGAAGGTACAGCTCGCGCAGAGATCTGATGATTTCTTACGACGACACAAC 890

Qy 186 TCAATTTGTCGACGACATGGGATTCCTTATGACGGTTGATGCCATGGCTGACGGGC 245

Db 891 TCAATTTGTCGACGACATGGGATTCCTTATGACGGTTGATGCCATGGCTGACGGGC 950

Qy 246 AGCCACTTTAAGATTGCACAAAGTCCCTGAGAGCGCTTCTCGGAGCTTAATAAGCCATTG 305

Db 951 AGCCACTTTAAGATTGCACAAAGTCCCTGAGAGCGCTTCTCGGAGCTTAATAAGCCATTG 1010

Qy 306 GGCAGAGATTTATGATATACAGGCTACCGTCTGTGACATGATGATGATGCGCTGGGA 365

Db 1011 GGCAGAGATTTATGATATACAGGCTACCGTCTGTGACATGATGATGATGCGCTGGGA 1070

Qy 366 GACAGAGGACACAGTCCGCGAGAAAGGCGCAATTTCTGCTGCGCAGCTCATGAGACAC 425

Db 1071 GACAGAGGACACAGTCCGCGAGAAAGGCGCAATTTCTGCTGCGCAGCTCATGAGACAC 1130

Qy 426 AGAGTCTTCCGCGCAGAGGCTGATGACAGCTGGCCACTAGCTCTTCAAGCACAG 485

Db 1131 AGAGTCTTCCGCGCAGAGGCTGATGACAGCTGGCCACTAGCTCTTCAAGCACAG 1190

Qy 486 AACGCCAAGGTGCGAGAGGATTCCTTACAGAGATTGTGAAGCTCTTCAAGTACGCG 545

Db 1191 AACGCCAAGGTGCGAGAGGATTCCTTACAGAGATTGTGAAGCTCTTCAAGTACGCG 1250

Qy 546 ACCCAGAGCTTATGTTGCGCTCTATATACCAAGTTGTGACCTTCTCGAGATCCC 605

Db 1251 ACCCAGAGCTTATGTTGCGCTCTATATACCAAGTTGTGACCTTCTCGAGATCCC 1310

Qy 606 ACGCTTAATGTAGGAGAGCGCGCATCCAAAGCTTGTGAATTTCAAGCATGTAGGG 665

Db 1311 ACGCTTAATGTAGGAGAGCGCGCATCCAAAGCTTGTGAATTTCAAGCATGTAGGG 1370

Qy 666 GATCGATTGGGCCACAGCTCGCTGATGAGAGATGTTCTGCTCGAAATTTGGCTATG 725

Db 1371 GATCGATTGGGCCACAGCTCGCTGATGAGAGATGTTCTGCTCGAAATTTGGCTATG 1430

QY 726 TTGGAGCAAAAGTTGCAACGAGTCAACAGAGGGTCTACTGCTTACCTTACGCTTAA 785
DB 1431 TTGGAGCAAAAGTTGCAACGAGTCAACAGAGGGTCTACTGCTTACCTTACGCTTAA 1490
QY 786 AACACGAATGCGCAATGAGAGGGCTTGGACGAGCCGCAAAATTTGGGTGAGGAGCGCA 845
DB 1491 AACACGAATGCGCAATGAGAGGGCTTGGACGAGCCGCAAAATTTGGGTGAGGAGCGCA 1550
QY 846 CCCACGAGATGATTAAGCGGCCTACTACACTCGGCCGTTTCTGTCATCACTGCGCCCAAA 905
DB 1551 CCCACGAGATGATTAAGCGGCCTACTACACTCGGCCGTTTCTGTCATCACTGCGCCCAAA 1610
QY 906 CCCAATGTGAACGATGTGACCGGTGATGCCGCGCGCTGAACCATGGAATCTTTGCAATCT 965
DB 1611 CCCAATGTGAACGATGTGACCGGTGATGCCGCGCGCTGAACCATGGAATCTTTGCAATCT 1670
QY 966 AGCTTTGAGGTGTCCTCCGCAATTGAACATCTTCCAGCTAAGACATGACATATCTAC 1025
DB 1671 AGCTTTGAGGTGTCCTCCGCAATTGAACATCTTCCAGCTAAGACATGACATATCTAC 1730
QY 1026 AAGCAAGTACTAGTATCATCAGTGTAAATAAAGCAGACTGGGAGAAAGTGTGATGCT 1085
DB 1731 AAGCAAGTACTAGTATCATCAGTGTAAATAAAGCAGACTGGGAGAAAGTGTGATGCT 1790
QY 1086 CTCGAAGATCAAGGCGATTTGCTCATTTCTCAGCTATCACTCAGCGCGAGTTTGTGCT 1145
DB 1791 CTCGAAGATCAAGGCGATTTGCTCATTTCTCAGCTATCACTCAGCGCGAGTTTGTGCT 1850
QY 1146 GTACAGCTAAAGAAATTTGCTTAAGCTTCGTGAGATATCTCAGAGAGAACTACAGATCA 1205
DB 1851 GTACAGCTAAAGAAATTTGCTTAAGCTTCGTGAGATATCTCAGAGAGAACTACAGATCA 1910
QY 1206 CAGGTATCCGCGAGGCGTGATCAACCATGCTTACATGCTTAAAGCGCTGAGAAATAA 1265
DB 1911 CAGGTATCCGCGAGGCGTGATCAACCATGCTTACATGCTTAAAGCGCTGAGAAATAA 1970
QY 1266 CTAGATGCTTCTGCTGAGCACTTTTGGAGCACTGTATTAATTAACAGAACAGCGCG 1325
DB 1971 CTAGATGCTTCTGCTGAGCACTTTTGGAGCACTGTATTAATTAACAGAACAGCGCG 2030
QY 1326 AAGGTCATTCGACCTCCGCTTCCAAATAGCTTGAAGTATATCATTAAGTATACATGCA 1385
DB 2031 AAGGTCATTCGACCTCCGCTTCCAAATAGCTTGAAGTATATCATTAAGTATACATGCA 2090
QY 1386 CCAAAAGTCTCAAGATCTTACAAGACACTCTGAATCAATCAAAAGTCGAGAGCATTAAG 1445
DB 2091 CCAAAAGTCTCAAGATCTTACAAGACACTCTGAATCAATCAAAAGTCGAGAGCATTAAG 2150
QY 1446 TTCACACTGTGTAGCTGATGCTGCTCTTTCGAGAGTGGCAGACGAGCGTTTGA 1505
DB 2151 TTCACACTGTGTAGCTGATGCTGCTCTTTCGAGAGTGGCAGACGAGCGTTTGA 2210
QY 1506 AGGAATGCCACCGTACTAGGAGACCTTAAATAAATTCATGCGGAGTCGACACTGCGAT 1565
DB 2211 AGGAATGCCACCGTACTAGGAGACCTTAAATAAATTCATGCGGAGTCGACACTGCGAT 2270
QY 1566 GCACGCCGCGCATTCAGATACCGCTTATTTGGGCTTTTCAGGCGCTCACTTTCAGAGCTGGCG 1625
DB 2271 GCACGCCGCGCATTCAGATACCGCTTATTTGGGCTTTTCAGGCGCTCACTTTCAGAGCTGGCG 2330
QY 1626 GATCAAAATATATAGAAATTTAAGCATTAAGCTGCCAGCGCATTTAGAAAGGAGAGAG 1685
DB 2331 GATCAAAATATATAGAAATTTAAGCATTAAGCTGCCAGCGCATTTAGAAAGGAGAGAG 2390
QY 1686 GCGCGCGGAG 1745
DB 2391 GCGCGCGGAG 2450
QY 1746 GTATCCCGGATTTGGCGGAACTCTGGAACCTTGCAAAAGCCGCTAGTATGAGATCC 1805
DB 2451 GTATCCCGGATTTGGCGGAACTCTGGAACCTTGCAAAAGCCGCTAGTATGAGATCC 2510
QY 1806 ATTTACGGGCTGACACTGCGGCTGCTCAACGAGCCAAAGTTAAGAGCGCAATATATCACTA 1865

DB 2511 ATTTACGGGCTGACACTGCGGCTGCTCAACGAGCCAAAGTTAAGAGCGCAATATCACTA 2570
QY 1866 TATTCAGGCAAAAGAAACCTTTAGAGCCTATATATTTCCAAACGAGCAATCGATGACAGA 1925
DB 2571 TATTCAGGCAAAAGAAACCTTTAGAGCCTATATATTTCCAAACGAGCAATCGATGACAGA 2630
QY 1926 GCAGAGCCAGTGTATCACTGCCAGACCTGCTGAAATTCCAATATAGCGGTGGCACCA 1985
DB 2631 GCAGAGCCAGTGTATCACTGCCAGACCTGCTGAAATTCCAATATAGCGGTGGCACCA 2690
QY 1986 GCTTAACAGCCGGGATTCGGTTTACACACAGGCCCGGGGAGCGGAGAGTATCTACATCA 2045
DB 2691 GCTTAACAGCCGGGATTCGGTTTACACACAGGCCCGGGGAGCGGAGAGTATCTACATCA 2750
QY 2046 CAGCAGAGATGAGATCCACCTCGCCAGACCAAAATGAGGAGATCAGATCGGTGATTT 2105
DB 2751 CAGCAGAGATGAGATCCACCTCGCCAGACCAAAATGAGGAGATCAGATCGGTGATTT 2810
QY 2106 GGTAAATTACTATCGTGGGGCCACTGGCGCATACCCCAAAAAGGCTCCGGAATACCCGA 2165
DB 2811 GGTAAATTACTATCGTGGGGCCACTGGCGCATACCCCAAAAAGGCTCCGGAATACCCGA 2870
QY 2166 AGCAGAGCCAGCTCTAGGAGAAACGAGTCAACTAGATCAGGTGTGTGCTTATGAAGCG 2225
DB 2871 AGCAGAGCCAGCTCTAGGAGAAACGAGTCAACTAGATCAGGTGTGTGCTTATGAAGCG 2930
QY 2226 AGTATGTACTCTACAGGTGTGGGGGCTGTGAGTACGCGCGAGAGAGAAACCCAGTAAGA 2285
DB 2931 AGTATGTACTCTACAGGTGTGGGGGCTGTGAGTACGCGCGAGAGAGAAACCCAGTAAGA 2990
QY 2286 CCATCGCGCGCGGACGACACTGTGCGGCAATCCCGTGAAGAGAAATATCATTTAGGCGTC 2345
DB 2991 CCATCGCGCGCGGACGACACTGTGCGGCAATCCCGTGAAGAGAAATATCATTTAGGCGTC 3050
QY 2346 GGAATATATGACAGCCGACGACTATGTTTCCGGGGAGCTACATGCCAGAGCGGGAATGCGG 2405
DB 3051 GGAATATATGACAGCCGACGACTATGTTTCCGGGGAGCTACATGCCAGAGCGGGAATGCGG 3110
QY 2406 ATGGGTAGAGAGCTCATGGAGAGTATGATGATCTGATGACATGACACTCGAGGCGAGTTCT 2465
DB 3111 ATGGGTAGAGAGCTCATGGAGAGTATGATGATGATGATGATGATGATGATGATGATGAT 3170
QY 2466 GTGTGTTCAGAACGATCTTTGACCTCAGCTACACTAGAGGTAAATTAATGAACTATCA 2525
DB 3171 GTGTGTTCAGAACGATCTTTGACCTCAGCTACACTAGAGGTAAATTAATGAACTATCA 3230
QY 2526 CTTAGCGGAGACCAACCCGCTTGGACTGAGAGCAACGAGCGGCGCCATTTGACGACATC 2585
DB 3231 CTTAGCGGAGACCAACCCGCTTGGACTGAGAGCAACGAGCGGCGCCATTTGACGACATC 3290
QY 2586 GAGAGATTTATTCAGTTCTGCGCATCGACGCAATTTGCTGAAAGAGAGAGAGAGAGAGAG 2645
DB 3291 GAGAGATTTATTCAGTTCTGCGCATCGACGCAATTTGCTGAAAGAGAGAGAGAGAGAG 3350
QY 2646 AGCCTCAACAGTATCTGCGCATGAGAAAGAGCTCACTCAGACAGCAACTTAATGCGTT 2705
DB 3351 AGCCTCAACAGTATCTGCGCATGAGAAAGAGCTCACTCAGACAGCAACTTAATGCGTT 3410
QY 2706 CTGACATGTTCCGCAAGATTTTATGATATCACACACAGAGGTGATCTCACTGTTTTTG 2765
DB 3411 CTGACATGTTCCGCAAGATTTTATGATATCACACACAGAGGTGATCTCACTGTTTTTG 3470
QY 2766 GACACGGGTACAGAACTGATCTGTTTCAATGAGAGAACTTACAGAAATGAGCTCTTCA 2825
DB 3471 GACACGGGTACAGAACTGATCTGTTTCAATGAGAGAACTTACAGAAATGAGCTCTTCA 3528
QY 2826 TCTTGTTCAGCGGTTTTCACAAATTTGGGACCGATCTCACTCAATTCATGATGACAGC 2885
DB 3529 TCTTGTTCAGCGGTTTTCACAAATTTGGGACCGATCTCACTCAATTCATGATGACAGC 3587
QY 2886 AAGATTTGAAGAGCTTACAGGTGTCTACAGAAATTTTCCCAAGCACTTTCAGCTTAA 2945

```
Db 3588 AAGATTGGAAGCGCTAGAGTGCTGCTCCAGAAATATTTCCCGCAGCTTGCACTTAA 3647
Qy 2946 GAACTATTTAGAAATCATATCGAATTTCTACTCAAAAGCCCACTAACAAAGCGCGCATCGCC 3005
Db 3648 GAACTATTTAGAAATCATATCGAATTTCTACTCAAAAGCCCACTAACAAAGCGCGCATCGCC 3707
Qy 3006 ATTTCGCGCTTCCCTTACGAGATCGGCGCAATATCTATGTAAGCAAGCGCACTTTCCGAC 3065
Db 3708 ATTTCGCGCTTCCCTTACGAGATCGGCGCAATATCTATGTAAGCAAGCGCACTTTCCGAC 3767
Qy 3066 GACCAAGAGCAGGCGCTGCGAGCGAGCGTCTTAAAGCTGAGCCAGCTGCGCGGAGTCAAG 3125
Db 3768 GACCAAGAGCAGGCGCTGCGAGCGAGCGTCTTAAAGCTGAGCCAGCTGCGCGGAGTCAAG 3827
Qy 3126 AAGTCGATGAGCTGCGCTCCAGGCGAGAGCTGCTAGTGGCCCTGTATTAACCTGAAT 3185
Db 3828 AAGTCGATGAGCTGCGCTCCAGGCGAGAGCTGCTAGTGGCCCTGTATTAACCTGAAT 3887
Qy 3186 ACCCGCGCAATGACCTTTTATCTGCGCCAGCTGCGCAAGATATATCAGGACTCTGCCGA 3245
Db 3888 ACCCGCGCAATGACCTTTTATCTGCGCCAGCTGCGCAAGATATATCAGGACTCTGCCGA 3947
Qy 3246 TCCGATCATCTGCGACATGAGGCGGCAAGCCAAAGTTGCAATTCGGGTGCCAATTCG 3305
Db 3948 TCCGATCATCTGCGACATGAGGCGGCAAGCCAAAGTTGCAATTCGGGTGCCAATTCG 4007
Qy 3306 CTTAGTAGCTCTCAATTGAGCAGTAGCAGTCCCAAGCCTTTGCAAGTCCCTCTGCGGT 3365
Db 4008 CTTAGTAGCTCTCAATTGAGCAGTAGCAGTCCCAAGCCTTTGCAAGTCCCTCTGCGGT 4067
Qy 3366 CCATTGCTGCTGCTTCAAGAGCCACCACTTAAGCATGCTCTTAATAGTCCAGCG 3425
Db 4068 CCATTGCTGCTGCTTCAAGAGCCACCACTTAAGCATGCTCTTAATAGTCCAGCG 4127
Qy 3426 TCCCGGAGCTCTCCGTGAGAGGAGGCTGCTTTCTCGGAGCTGAGCACTTAGAGAC 3485
Db 4128 TCCCGGAGCTCTCCGTGAGAGGAGGCTGCTTTCTCGGAGCTGAGCACTTAGAGAC 4187
Qy 3486 AACATTGAGAAAGCTGCGAGAGAGATCCGCGCACTGCTGCGGCTCAGTATCAGACGCG 3545
Db 4188 AACATTGAGAAAGCTGCGAGAGAGATCCGCGCACTGCTGCGGCTCAGTATCAGACGCG 4247
Qy 3546 CTGGCGCCCAATGCTTCAATGAGACATTTGCAATGATCAGATCAGGCGCAACAGATTCG 3605
Db 4248 CTGGCGCCCAATGCTTCAATGAGACATTTGCAATGATCAGATCAGGCGCAACAGATTCG 4307
Qy 3606 TGTGATCCCTGCTTCCAACTCCAAAGCGCAATGCTGCGGCACTAACCTAGTCAAT 3665
Db 4308 TGTGATCCCTGCTTCCAACTCCAAAGCGCAATGCTGCGGCACTAACCTAGTCAAT 4367
Qy 3666 ACACCTGAGTACAGCAACATGAGGCTGATATCTGAGCGGGAAGAGCACTCAGAGC 3725
Db 4368 ACACCTGAGTACAGCAACATGAGGCTGATATCTGAGCGGGAAGAGCACTCAGAGC 4427
Qy 3726 GCCAAGTACCACTGAGATGCCAAGGTGATCAAGCTCTGATTAATATGCTGAAT 3785
Db 4428 GCCAAGTACCACTGAGATGCCAAGGTGATCAAGCTCTGATTAATATGCTGAAT 4487
Qy 3786 GGAAGCTGATATCTGCGCAGCACTTATGAGAGCGCAAGTGTGCGTGTGCTTGAAG 3845
Db 4488 GGAAGCTGATATCTGCGCAGCACTTATGAGAGCGCAAGTGTGCGTGTGCTTGAAG 4547
Qy 3846 CTAAACAAGATTCAGCCCGTGAAGTGTGCTTCAAGAGTCACTTAACCTGGGATTTGC 3905
Db 4548 CTAAACAAGATTCAGCCCGTGAAGTGTGCTTCAAGAGTCACTTAACCTGGGATTTGC 4607
Qy 3906 ATCAAGGCGGAACTGTGAGCTGCCAATAGACATTTAGATGATGATGCAATGCAATGCTG 3965
Db 4608 ATCAAGGCGGAACTGTGAGCTGCCAATAGACATTTAGATGATGATGCAATGCAATGCTG 4667
Qy 3966 CTTAACATTTCTGAGGCGGAGCATACGAGCTGTGATCTGCTGCGCTGACGCTGCTAGT 4025
Db 4668 CTTAACATTTCTGAGGCGGAGCATACGAGCTGTGATCTGCTGCGCTGACGCTGCTAGT 4727

Qy 4026 AAGATTATGAGAGCAACAAATGCGTCAAACTGATGATCACTTTCTAGAGCTGATTTTG 4085
Db 4728 AAGATTATGAGAGCAACAAATGCGTCAAACTGATGATCACTTTCTAGAGCTGATTTTG 4787
Qy 4086 CTGAAGATATTCAGTGTATATCAACAGCAAGAGGCTTTGCGGGATATGATCTGATG 4145
Db 4788 CTGAAGATATTCAGTGTATATCAACAGCAAGAGGCTTTGCGGGATATGATCTGATG 4847
Qy 4146 ATACCAAGATATGACCACTCTGCTGATGATGTCATCAATGATGTCATTCGGA 4205
Db 4848 ATACCAAGATATGACCACTCTGCTGATGATGTCATCAATGATGTCATTCGGA 4907
Qy 4206 ATTGCGAGGTTGATTTCCCAAAATTTGCGCCATCAAGATCTCTGGAATGAC 4265
Db 4908 ATTGCGAGGTTGATTTCCCAAAATTTGCGCCATCAAGATCTCTGGAATGAC 4967
Qy 4266 GAACACATGTTGAGATCAACAGCGCCACTTGAATGATGTTGCCAATCTGCG 4325
Db 4968 GAACACATGTTGAGATCAACAGCGCCACTTGAATGATGTTGCCAATCTGCG 5027
Qy 4326 CGATCGCGGAGCAACAGCAATGATGTTGCAAAAGTGGGTCTTGCATGCTCAAG 4385
Db 5028 CGATCGCGGAGCAACAGCAATGATGTTGCAAAAGTGGGTCTTGCATGCTCAAG 5087
Qy 4386 CTGATCTTTGTTCTGGCGAGAGAAAGTCAAGCCGAAGCTGTAGTCTAAATCCAGC 4445
Db 5088 CTGATCTTTGTTCTGGCGAGAGAAAGTCAAGCCGAAGCTGTAGTCTAAATCCAGC 5147
Qy 4446 AAGTTAGGCTCTCAACGTGTACATGAGAGAGCGGAACTGATCATGAGGAGAG 4505
Db 5148 AAGTTAGGCTCTCAACGTGTACATGAGAGAGCGGAACTGATCATGAGGAGAG 5207
Qy 4506 AGCTCTCAAAAGAACTCTCCGCGGCACTGCTGCTCAATGTCGGAGCCCTTAATAGA 4565
Db 5208 AGCTCTCAAAAGAACTCTCCGCGGCACTGCTGCTCAATGTCGGAGCCCTTAATAGA 5267
Qy 4566 TTCTCTGC--TCGTGCGACCAACAAACAGACAGACGCGTGTCTTCCCTGCGCTGAGA 4623
Db 5268 TTCTCTGATACAGACCAACAAACAGACAGACGCGTGTCTTCCCTGCGCTGAGA 5327
Qy 4624 AGGAAGTGAAGAGGCGGCACTTAATTAATTAATTAATTAATTAATTAATTAATTA 4683
Db 5328 AGGAAGTGAAGAGGCGGCACTTAATTAATTAATTAATTAATTAATTAATTAATTA 5387
Qy 4684 CTATATTTAACGATCATCCGTTACTGTGTACTGTGATGATGATATTCACAGATG 4743
Db 5388 CTATATTTAACGATCATCCGTTACTGTGTACTGTGATGATGATATTCACAGATG 5447
Qy 4744 CAATGCGGCGGCAAGAAAGTCAAGGCGCATGCGCCATCTAATGTAGAGAGAAATCTA 4803
Db 5448 CAATGCGGCGGCAAGAAAGTCAAGGCGCATGCGCCATCTAATGTAGAGAGAAATCTA 5507
Qy 4804 TTATATCATTAACGGAACCTTAAGCGAGTAATCCGCAAGTGTACTTAATTTAGCCA 4863
Db 5508 TTATATCATTAACGGAACCTTAAGCGAGTAATCCGCAAGTGTACTTAATTTAGCCA 5567
Qy 4864 ACTCCGAATCTCTCATCTTCTCATCATGATGTGACGCTTTTATACAGCTCTTAATTA 4923
Db 5568 ACTCCGAATCTCTCATCTTCTCATCATGATGTGACGCTTTTATACAGCTCTTAATTA 5627
Qy 4924 ACCGTGATCTGTATTAATTTGTAAGCTATATACCGCTCTTATGGAACCTCAGACGTG 4983
Db 5628 ACCGTGATCTGTATTAATTTGTAAGCTATATACCGCTCTTATGGAACCTCAGACGTG 5687
Qy 4984 GCTTTTCTACCTTTGTTTACGCGCACTTGTATCTTATTTGCTTCCGCTCAACTAAT 5043
Db 5688 GCTTTTCTACCTTTGTTTACGCGCACTTGTATCTTATTTGCTTCCGCTCAACTAAT 5747
Qy 5044 TCTAATTTACTATTAATTAACGATTTGCTCTTACAACTGAACATTTTGTATTAAT 5103
Db 5748 TCTAATTTACTATTAATTAACGATTTGCTCTTACAACTGAACATTTTGTATTAAT 5807
```


QY 5104 TAAATACACAGCCCAACAGAGATTAAATCAATPAAAAA 5145
 DB 5808 TAAATACACAGCCCAACAGAGATTAAATCAATPAAAAA 5849

RESULT 4
 AF250842 5938 bp mRNA linear INV 04-AUG-2000
 LOCUS Drosophila melanogaster multiple asters (Mast) mRNA, complete cds.
 DEFINITION AF250842
 ACCESSION AF250842.1 GI:7650478
 VERSION
 KEYWORDS
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 5938)
 Lemos, C.L., Sampaio, P., Maisto, H., Costa, M., Omei, Yanchuk, L.V.,
 Liberal, V., and Sunkel, C.E.
 Mast, a conserved microtubule-associated protein required for
 bipolar mitotic spindle organization
 EMBO J 19 (14), 3668-3682 (2000)
 2 (bases 1 to 5938)
 Maisto, H., Lemos, C.L., Sampaio, P., and Sunkel, C.E.
 Direct Submission
 Submitted (30-MAR-2000) Genetics Molecular da Mitose, Instituto de
 Biologia Molecular e Celular, Rua Campo Alegre, 823, Porto
 4150-180, Portugal
 Location/Qualifiers
 1. 5938
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /chromosome="3"
 /map="3L: 78C1-C2"
 /clone="LD11488"
 /dev_stage="0-24 hours embryo"
 1. 5938
 /gene="Mast"
 /gene="Mast"
 /note="Mast; microtubule-associated protein; essential for
 centrosome separation and organization of mitotic
 apparatus"
 /codon_start=-1
 /product="multiple asters"
 /protein_id="AA06060.1"
 /db_xref="GI:7650478"
 /translation="MAVRKPSDLDFIQMPKADMRKVQLAEDLVPLSDTNSIVC
 TMGFLIDGMPWLTGSHFKIAOKSLAESELIKRGSFNATATVLPVHIDRIGDS
 RDTVRKAQLLRDLMEHRYLPQALIDKLATCFPKGNKARBEELQITVNAHYG
 TQGLSVRYTIPYCALGDPYVVRRAALQTVLEIKYKQDRLPRDMDVPASKL
 AMLEQKPDVQKQGLLPSALKNTNGGGLDEADNIGLRPERTKRIKPLSAYSS
 LRPEKPNVDTGAGAVTWESFESEFVQPLNI FPAKMDIYKQVLIVISDKNAD
 EKVDLKKIRALLIISYHQPQFVAQKEISLSTVDILKELRQVIREACITAY
 MSKTLNKLDAPFMSIIEHLINLONSARKIASATIALKYIIKYHAPRLKIYDT
 LNSKSKDIRSTICELMWLFEEMOTKALERNATVLRDLTKKISIGADDDARHSRY
 YAFRRHPBELAQIYGTIDIAQRALREBEGGGGTGTGTGTPERRRYSRGR
 TPTGLKPTPSMRISAVDTAAQAKRVAQVLYISRKRPVPPNSNASTMAGN
 GSLPRRLNSNGSPTATPGVTPPRAGVSOQPSRSTSPSKLRDYGGIGN
 YRGATGAI PKKASGIPRSTASRETSPTSGGLMKRISVTSAGASRTPERNMVR
 PSAARALAOISRAEHTLGVDDGPDYVSGDYMRSGMRGKLMGRDESDIDSEA
 SVCSRSPDSVYTRGNKSYLSGSHTLDMSTORAPPDITIIIOFCAJHMSRX
 DGLISTVYLABSKETCOOLKCVLDMFPMKPMDTKTVYSFLDTPTLLVHANEL
 HEMFLTLRLFKLGTDLNMSWKIMTKLOVYHAYFTQLOKELFRIISDTYTP
 TYKTRIALIRFTDLANTYCKSSDPSPDSQACERTVLKLAQAAQKMELSQKRS
 CVALYNLNTPQWTLTLADLPKYQDSASCSHMRROSQNSGANSPPSSPLSS
 SPKLPSPSVGPASIOSHHOLISITSPRSRQSVVEIDFLFSLDIQHNOKTSE
 ETRHCGGQYOTLAPNGFNHGLIYHDOGGQSCASLSNSKQSSANTQNTSPSA
 TMRLDUBERTTQNAKSPETDAKVITTVISMAENBELILIASIMSEVRYALTTTK
 DQVEILLQTSINLIGICIKGNCELPKFKRSIMRWLNLILEHTDVVYIAGLHVISK

ORIGIN

Query Match 98.3%; Score 5055.6; DB 2; Length 5938;
 Best Local Similarity 99.3%; Pred. No. 0; Mismatches 29; Indels 5; Gaps 3;
 Matches 5108; Conservative 0;

IRMSKRNHMMHLELILKLIQCYHSEKALDIDSMRIAPSLDLSINLVN
 VIATGEFTNLCAIKILEVTEHHGSEITDHLIDVFNLRASDDTSMRKAAVPC
 IYKVYFVIGEEKVPRKLSVNLNPSKVRLLNVIYKORNCISGGGSSTKNSAASS"

QY 6 CGGACGAGCCGCTCCGAATTTGTTGGCGAGAACGAGACGATGACATTGCAAG 65
 DB 710 CAGCGACCCGCGATTGGAATAGTTGCGGAGAACGAGACGATGACATTGCAAG 769
 QY 66 ATGCGCTTACGAGACCCGAGCAGCAGCTGATGCTTTATTCAGCAATGCCAGCGGAC 125
 DB 770 ATGCGCTTACGAGACCCGAGCAGCAGCTGATGCTTTATTCAGCAATGCCAGCGGAC 829
 QY 126 ATGCGTGAAGATGACAGCTGCGCGAGGATCTGGGACATTCCTTAGGCGAGCACAAA 185
 DB 830 ATGCGTGAAGATGACAGCTGCGCGAGGATCTGGGACATTCCTTAGGCGAGCACAAA 889
 QY 186 TCAATTTGCTGACGAGCAGATGGGATTCCTTATTTGACGCGTTGATGCAATGGCTGACG 245
 DB 890 TCAATTTGCTGACGAGCAGATGGGATTCCTTATTTGACGCGTTGATGCAATGGCTGACG 949
 QY 246 AGCCACTTTAAGATTCGACCAAAAGTCCCTGGAGGCGCTTCTGGAGCTATTAAGGAT 305
 DB 950 AGCCACTTTAAGATTCGACCAAAAGTCCCTGGAGGCGCTTCTGGAGCTATTAAGGAT 1009
 QY 306 GGGACGAGTTTAAAGCATACAGCGCTACCGTTCTGACCAATGATGATCGGCTGGGA 365
 DB 1010 GGGACGAGTTTAAAGCATACAGCGCTACCGTTCTGACCAATGATGATCGGCTGGGA 1069
 QY 366 GACAGCAGGAGCACAGTCCGCGAGAGAGCGCAACTTCTGCTGCGGACCTCATGAGAC 425
 DB 1070 GACAGCAGGAGCACAGTCCGCGAGAGAGCGCAACTTCTGCTGCGGACCTCATGAGAC 1129
 QY 426 AGAGTCTTCGCGCCCGACCGCTGATTCGACAAAGCTGGCCACTGCTTCAAGACAA 485
 DB 1130 AGAGTCTTCGCGCCCGACCGCTGATTCGACAAAGCTGGCCACTGCTTCAAGACAA 1189
 QY 486 AACGCAAGAGGCGCGAGAGAGTTCCTTACAGCATTTGAGACGCTCTTATGATAGTAC 545
 DB 1190 AACGCAAGAGGCGCGAGAGAGTTCCTTACAGCATTTGAGACGCTCTTATGATAGTAC 1249
 QY 546 ACCGACAGCTTATGTTGCTGCTTATATACCAAGCTTTGCACTTCTGAGATGCC 605
 DB 1250 ACCGACAGCTTATGTTGCTGCTTATATACCAAGCTTTGCACTTCTGAGATGCC 1309
 QY 606 ACAATTAATGAGGAGGAGCGGCGCATCCAAAGCTAGTGGAAATCTACAGCATGTAGG 665
 DB 1310 ACAATTAATGAGGAGGAGCGGCGCATCCAAAGCTAGTGGAAATCTACAGCATGTAGG 1369
 QY 666 GATGATGCGCGCGAGCCTCGCTGATGAGCAATGTTCTGCTCGAAATTTGGCTATG 725
 DB 1370 GATGATGCGCGCGAGCCTCGCTGATGAGCAATGTTCTGCTCGAAATTTGGCTATG 1429
 QY 726 TTGGAGCAAAAGTTGACCAAGTCAAAAGAGGCTTACTGCTTCACTTCACTTAA 785
 DB 1430 TTGGAGCAAAAGTTGACCAAGTCAAAAGAGGCTTACTGCTTCACTTCACTTAA 1489
 QY 786 AACAGCAATGCGCAATGAGTGGGCTTGAACGAGGCGGCAATATTTGGTGGAGGAGCA 845
 DB 1490 AACAGCAATGCGCAATGAGTGGGCTTGAACGAGGCGGCAATATTTGGTGGAGGAGCA 1549
 QY 846 CCCACAGAGATTAAGGCGGCACTACACTGCGCGCTTCTCATCTAGCGCCCAAA 905
 DB 1550 CCCACAGAGATTAAGGCGGCACTACACTGCGCGCTTCTCATCTAGCGCCCAAA 1609
 QY 906 CCCAATGTAACGATGTGACCGGTGATCGCGGCGGCTTACCATGATCTTTGCAATCT 965
 DB 1610 CCCAATGTAACGATGTGACCGGTGATCGCGGCGGCTTACCATGATCTTTGCAATCT 1669

QY 966 AGCTTTGAGTGTGCTCCGCAATGAACTCTTCCAGCTTAAAGACATGACGATATCTAC 1025
DB 1670 AGCTTTGAGTGTGCTCCGCAATGAACTCTTCCAGCTTAAAGACATGACGATATCTAC 1729
QY 1026 AAGCAAGTCTAGTGTATCATCATGATGATTAATAAAGCAGACTGGGAAAGCGTGTGATGCT 1085
DB 1730 AAGCAAGTCTAGTGTATCATCATGATGATTAATAAAGCAGACTGGGAAAGCGTGTGATGCT 1789
QY 1086 CTCGAAGAATCATGAGGCAATGCTCATTTCTAGCTATCACACTGAGCGGATTTGTGCT 1145
DB 1790 CTCGAAGAATCATGAGGCAATGCTCATTTCTAGCTATCACACTGAGCGGATTTGTGCT 1849
QY 1146 GTACAGCTAAAGGAATTTGCTTAAAGCTTGTGAGCACTCTCAAGAGAACTACGATCA 1205
DB 1850 GTACAGCTAAAGGAATTTGCTTAAAGCTTGTGAGCACTCTCAAGAGAACTACGATCA 1909
QY 1206 CAGGTATCCGCGAGGCTGATCACTATGCTTCAAGCTGAGCAAGCTGAGAAATTA 1265
DB 1910 CAGGTATCCGCGAGGCTGATCACTATGCTTCAAGCTGAGCAAGCTGAGAAATTA 1969
QY 1266 CTAGATGCTTCTGCTGAGACATTTTGGAGCACTGATTAATTAATACAGACGCG 1325
DB 1970 CTAGATGCTTCTGCTGAGACATTTTGGAGCACTGATTAATTAATACAGACGCG 2029
QY 1326 AAGGTATGTCATCCGCTTCCAGAAATAGCTCTGAAGTATCATTAATATACATGCA 1385
DB 2030 AAGGTATGTCATCCGCTTCCAGAAATAGCTCTGAAGTATCATTAATATACATGCA 2089
QY 1386 CCAAGCTGCTCAAGATCTTACACAGACACTCTGATATCAAAAGTCAAGAGCATTAAG 1445
DB 2090 CCAAGCTGCTCAAGATCTTACACAGACACTCTGATATCAAAAGTCAAGAGCATTAAG 2149
QY 1446 TCCACACTGCTGAGTGTGATGCTGCTTCCAGAGAGTGCAGACGAAGGCTTGA 1505
DB 2150 TCCACACTGCTGAGTGTGATGCTGCTTCCAGAGAGTGCAGACGAAGGCTTGA 2209
QY 1506 AGGAATGCCACCGTACTTAAGGGAACCTTAAAAAATCATTGGCGATGCGAT 1565
DB 2210 AGGAATGCCACCGTACTTAAGGGAACCTTAAAAAATCATTGGCGATGCGAT 2269
QY 1566 GCAAGCCGCTTCCAGATAGCGCTTATTTGGCTTTCAGGCTCACTTTCCAGAGCTGGCG 1625
DB 2270 GCAAGCCGCTTCCAGATAGCGCTTATTTGGCTTTCAGGCTCACTTTCCAGAGCTGGCG 2329
QY 1626 GATCAATATATGGAACATTAAGACATAGCTGCCAGCGGCTTGAAGAGGAGACGAG 1685
DB 2330 GATCAATATATGGAACATTAAGACATAGCTGCCAGCGGCTTGAAGAGGAGACGAG 2389
QY 1686 GCGGCGGAGAGAGAGACCTGTACTGGACTGGGACTGCACTGCACTGAACGAGCGACT 1745
DB 2390 GCGGCGGAGAGAGAGACCTGTACTGGACTGGGACTGCACTGCACTGAACGAGCGACT 2449
QY 1746 GTATCCGCTTGGCCGAACACTGGAACCTCTGAAGAGCCGCTTATGATGATCC 1805
DB 2450 GTATCCGCTTGGCCGAACACTGGAACCTCTGAAGAGCCGCTTATGATGATCC 2509
QY 1806 ATTTCAAGCGGTGACACTGGGCTGCTCAAGAGCCAAAGTTAGAGGCAATATACCTA 1865
DB 2510 ATTTCAAGCGGTGACACTGGGCTGCTCAAGAGCCAAAGTTAGAGGCAATATACCTA 2569
QY 1866 TATTTCCAGGAAAGAAACCTTTAGACCTTAATTAATTCACACAGGATGATGACAGA 1925
DB 2570 TATTTCCAGGAAAGAAACCTTTAGACCTTAATTAATTCACACAGGATGATGACAGA 2629
QY 1926 GCAGCAGCCAGTGTATCATCTGCCAGACTGCTGCTGAATTCATAGCGGTGGACCA 1985
DB 2630 GCAGCAGCCAGTGTATCATCTGCCAGACTGCTGCTGAATTCATAGCGGTGGACCA 2689
QY 1986 GCTTCAACCGCGGATGCTTACACACGCGCCCGGGGAGCGGAGAGTATCTCAGTCA 2045
DB 2690 GCTTCAACCGCGGATGCTTACACACGCGCCCGGGGAGCGGAGAGTATCTCAGTCA 2749
QY 2046 CAGCAGGATGAGATCCACTCGCCAAAGCAAAACTGAAGGATCAATGCTGATTT 2105

DB 2750 CAGCAGGATGAGATCCACTCGCCAAAGCAAAACTGAAGGATCAATGAGTGTATTT 2809
QY 2106 GGTAAATTAATCTATGCTGGGCGCATGCGCCCATATCCCAAAAGGCTCCGGAATACCCCA 2165
DB 2810 GGTAAATTAATCTATGCTGGGCGCATGCGCCCATATCCCAAAAGGCTCCGGAATACCCCA 2869
QY 2166 AGCAGCAGCTCTAAGGAAAGAGAGTCCAACTAGTCAAGTGTGGCTGATGAAGCGC 2225
DB 2870 AGCAGCAGCTCTAAGGAAAGAGAGTCCAACTAGTCAAGTGTGGCTGATGAAGCGC 2929
QY 2226 AGTATGATCTTCAAGTGTGGGCTCTGCACTGACCCGAGAGAAACAACCAATGAGA 2285
DB 2930 AGTATGATCTTCAAGTGTGGGCTCTGCACTGACCCGAGAGAAACAACCAATGAGA 2989
QY 2286 CCATCGGCGCGGCAAGACTGTGGCGCAATCCGCTGAAGCAAAATATAGGCGTTC 2345
DB 2990 CCATCGGCGCGGCAAGACTGTGGCGCAATCCGCTGAAGCAAAATATAGGCGTTC 3049
QY 2346 GGAATGATGAGACAGCAGACTATGTTTCCGGGAGCTACATGCGGACGCGGGAATGCGG 2405
DB 3050 GGAATGATGAGACAGCAGACTATGTTTCCGGGAGCTACATGCGGACGCGGGAATGCGG 3109
QY 2406 ATGGGTAGGAAGCTGATGGGAGCTGATGAGTGAATGACATGCACTCCGAGGCAATGCT 2465
DB 3110 ATGGGTAGGAAGCTGATGGGAGCTGATGAGTGAATGACATGCACTCCGAGGCAATGCT 3169
QY 2466 GTGTGTTGAGAAAGATCTTTGCACTCCAGCTCACTAGAGGTAAATATGAACTTACTCA 2525
DB 3170 GTGTGTTGAGAAAGATCTTTGCACTCCAGCTCACTAGAGGTAAATATGAACTTACTCA 3229
QY 2526 CTTAGCGGAGGCAACATCCGCTTGGACTGAGAGACGCAAGCGGCGCATTTGACGACATC 2585
DB 3230 CTTAGCGGAGGCAACATCCGCTTGGACTGAGAGACGCAAGCGGCGCATTTGACGACATC 3289
QY 2586 GAGACGATATATGATCTGGGCAATGAGCACTGAGTGTGAAAGGAAGATGGCTGATC 2645
DB 3290 GAGACGATATATGATCTGGGCAATGAGCACTGAGTGTGAAAGGAAGATGGCTGATC 3349
QY 2646 AGCTTCAACAGATATGAGCCGATGGAAGAGAGCTCACTCAGACGAACCTTAATGCTT 2705
DB 3350 AGCTTCAACAGATATGAGCCGATGGAAGAGAGCTCACTCAGACGAACCTTAATGCTT 3409
QY 2706 CTGACATGTTCCGCAAGATGTTTATGATATACACACCAAGGTGATCACTGTTTGG 2765
DB 3410 CTGACATGTTCCGCAAGATGTTTATGATATACACACCAAGGTGATCACTGTTTGG 3469
QY 2766 GACACGTCACGGAATGATCTGCTGCTTCAATGGGAAGAACTTCAAGAAATGCTCTTCA 2825
DB 3470 GACACGTCACGGAATGATCTGCTGCTTCAATGGGAAGAACTTCAAGAAATGCTCTTCA 3527
QY 2826 TCTTGTGTAAGCGGCTGTTTCAACAAATTTGGGACCCGATCTCAATTCGATGACAGC 2885
DB 3528 TCTTGTGTAAGCGGCTGTTTCAACAAATTTGGGACCCGATCTCAATTCGATGACAGC 3586
QY 2886 AAGATTTGGAAGCGCTACAGGTGTCCAGCAATATTTCCCGCAGGATGACGCTTAA 2945
DB 3587 AAGATTTGGAAGCGCTACAGGTGTGTCCAGCAATATTTCCCGCAGGATGACGCTTAA 3646
QY 2946 GAATCTTTTGAATCAATTCGATTTCTACTCAAAAGCCAACTCAAGAGACGCGCATGCGC 3005
DB 3647 GAATCTTTTGAATCAATTCGATTTCTACTCAAAAGCCAACTCAAGAGACGCGCATGCGC 3706
QY 3006 ATCTGCGCTTCTTACGATCTGAGCAATATACATCTGTAAGAGAGCACTTTCCAGC 3065
DB 3707 ATCTGCGCTTCTTACGATCTGAGCAATATACATCTGTAAGAGAGCACTTTCCAGC 3766
QY 3066 GACCAAGCCAGGCTTGCAGAGGAGAGTCTTAAAGCTGGCCCAAGTGTGGGAGATCAG 3125
DB 3767 GACCAAGCCAGGCTTGCAGAGGAGAGTCTTAAAGCTGGCCCAAGTGTGGGAGATCAG 3826
QY 3126 AATTCATGAGAGTGGGCTCCCAAGGACAGAGGTGCTAGTGGCCCTGTAATCTGTAAT 3185

Db 3827 AAGTCATGAGCTGCGCTCCAGGCCAGAGCTGCTAGTGGCCCTGTATACCTGAAT 3886
 Qy 3186 ACCCGCAATGACCCCTTTTACTGCGCCGACCTGCCAAGATATATCAGACTCTGCCGA 3245
 Db 3887 ACCCGCAATGACCCCTTTTACTGCGCCGACCTGCCAAGATATATCAGACTCTGCCGA 3946
 Qy 3246 TCTGTCATCCATTTGGACATAGGGGCGCAAGCCAAAGTTGCCAATTTGGGTCCTCAATTCG 3305
 Db 3947 TCTGTCATCCATTTGGACATAGGGGCGCAAGCCAAAGTTGCCAATTTGGGTCCTCAATTCG 4006
 Qy 3306 CCTAGTAGCTCTCATTTGAGAGATAGACGTCCCAAGCTTTGCAAGCTTCCTGTGGGT 3365
 Db 4007 CCTAGTAGCTCTCATTTGAGAGATAGACGTCCCAAGCTTTGCAAGCTTCCTGTGGGT 4066
 Qy 3366 CCATTTGCTGCTTTCAGAGCCACCACTTATGATCAGCTCTACTAGTCCAGC 3425
 Db 4067 CCATTTGCTGCTTTCAGAGCCACCACTTATGATCAGCTCTACTAGTCCAGC 4126
 Qy 3426 TCCCGGAGTCTTCCGTGGAGAGAGAGTCTTTTCTTCTGGAGCTGGACATTCAGAC 3485
 Db 4127 TCCCGGAGTCTTCCGTGGAGAGAGAGTCTTTTCTTCTGGAGCTGGACATTCAGAC 4186
 Qy 3486 AACATTGAGAGAGTGGAGAGAGATCCGGACCTTGGCGGTCACTCCAGAGCGG 3545
 Db 4187 AACATTGAGAGAGTGGAGAGAGATCCGGACCTTGGCGGTCACTCCAGAGCGG 4246
 Qy 3546 CTGGCGCCCAATGCTTCAATGACACTTGCATGATCAGATCAGGCGCCAAAGATTCG 3605
 Db 4247 CTGGCGCCCAATGCTTCAATGACACTTGCATGATCAGATCAGGCGCCAAAGATTCG 4306
 Qy 3606 TGTGCAATCCCTGCTTTTCCAACTCCAAAGCGAATGCTGGCGCAACTTACCAGTCAAT 3665
 Db 4307 TGTGCAATCCCTGCTTTTCCAACTCCAAAGCGAATGCTGGCGCAACTTACCAGTCAAT 4366
 Qy 3666 AACCTGAGTCAGCAACATGAGGCTGATATATGAGGCGGGAAGAGCACTCAGAAC 3725
 Db 4367 AACCTGAGTCAGCAACATGAGGCTGATATATGAGGCGGGAAGAGCACTCAGAAC 4426
 Qy 3726 GCCAAGTCACCACTGACGATGCCAAGGTATCACTGCTTCAATTAATTTGGTGAAT 3785
 Db 4427 GCCAAGTCACCACTGACGATGCCAAGGTATCACTGCTTCAATTAATTTGGTGAAT 4486
 Qy 3786 GGAGAGCTGATCTGCGGACAGCACTGATGAGAGCGCAAGTGGTGGCTTGAACG 3845
 Db 4487 GGAGAGCTGATCTGCGGACAGCACTGATGAGAGCGCAAGTGGTGGCTTGAACG 4546
 Qy 3846 CTACAAAGATCAGCCCGTGGAGTGTTCAGAGCTCACTACTCACTGCGGGAATTCG 3905
 Db 4547 CTACAAAGATCAGCCCGTGGAGTGTTCAGAGCTCACTACTCACTGCGGGAATTCG 4606
 Qy 3906 ATCAAGGGCGGAACTGTGAGCTGCCAATTAAGCACTTTAGATGATCATGCGAATGCTG 3965
 Db 4607 ATCAAGGGCGGAACTGTGAGCTGCCAATTAAGCACTTTAGATGATCATGCGAATGCTG 4666
 Qy 3966 CTTTAACATTTCTGAGGCGGAGCAATAGGACGTGTCTATCCGTGCTGCAAGCTGCTAGT 4025
 Db 4667 CTTTAACATTTCTGAGGCGGAGCAATAGGACGTGTCTATCCGTGCTGCAAGCTGCTAGT 4726
 Qy 4026 AAGATTATGAGAGCAAAATGCTGACAACTGAGTGCATTTCTAGAGCTGATTTTG 4085
 Db 4727 AAGATTATGAGAGCAAAATGCTGACAACTGAGTGCATTTCTAGAGCTGATTTTG 4786
 Qy 4086 CTGAAGATCATCAGTCTATCAACAGAGAGAGGCTTTGCGGATATGACTGATG 4145
 Db 4787 CTGAAGATCATCAGTCTATCAACAGAGAGAGGCTTTGCGGATATGACTGATG 4846
 Qy 4146 ATACCAAGATAGACATCTTGGCTCTGATATGTCATCAATTTGCAATTCGGTA 4205
 Db 4847 ATACCAAGATAGACATCTTGGCTCTGATATGTCATCAATTTGCAATTCGGTA 4906
 Qy 4206 ATTCGAGCGGAGTGTTCACAAATTTGTGCGCCATCAAGATCTGCTGAGAGTACC 4265
 Db 4907 ATTCGAGCGGAGTGTTCACAAATTTGTGCGCCATCAAGATCTGCTGAGAGTACC 4966

Qy 4266 GAAACACATGTTTCGAGATCAACAGCGCCCACTTGGACATGTGTTCGCCAATTCGCG 4325
 Db 4967 GAAACACATGTTTCGAGATCAACAGCGCCCACTTGGACATGTGTTCGCCAATTCGCG 5026
 Qy 4326 CGATCGCGGAGCAACCGCATGATGTGTTCCGAAAGCTGGGTCTTTTGCATTCGTAAG 4385
 Db 5027 CGATCGCGGAGCAACCGCATGATGTGTTCCGAAAGCTGGGTCTTTTGCATTCGTAAG 5086
 Qy 4386 CTGTACTTTGTTTGGGCGAGGAGAGTCAAGCCGAGCTGTCACTGCTAAATCCAGC 4445
 Db 5087 CTGTACTTTGTTTGGGCGAGGAGAGTCAAGCCGAGCTGTCACTGCTAAATCCAGC 5146
 Qy 4446 AAGTTAGGCTCTCAACGTGTATCATGAGAGAGAGGAACTGCATCAGTGGGAGGA 4505
 Db 5147 AAGTTAGGCTCTCAACGTGTATCATGAGAGAGAGGAACTGCATCAGTGGGAGGA 5206
 Qy 4506 AGCTCTACAAAGAACTCTCCGCGGATGTGTGTATGATTTGGGAGAGCCCTTAATAGA 4565
 Db 5207 AGCTCTACAAAGAACTCTCCGCGGATGTGTGTATGATTTGGGAGAGCCCTTAATAGA 5266
 Qy 4566 TTCCTGCT--TCGCGCCACCAAAACAAGACAGAGCGGCTTCCCTCGGCTGAGA 4623
 Db 5267 TTCCTGCTGACGACACCAACAAAGACAGAGCGGCTTCCCTCGGCTGAGA 5326
 Qy 4624 AGGAAGTGAAGAGCGCGGACATTAATTAATTAATTAATTAATTAATTAATTAATTA 4683
 Db 5327 AGGAAGTGAAGAGCGCGGACATTAATTAATTAATTAATTAATTAATTAATTAATTA 5386
 Qy 4684 CTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4743
 Db 5387 CTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5446
 Qy 4744 CAAATCGGCGCGGAAAGAAAGTCAAGGCGCATCGCCATCTAATGTGAGCAGAAAACTA 4803
 Db 5447 CAAATCGGCGCGGAAAGAAAGTCAAGGCGCATCGCCATCTAATGTGAGCAGAAAACTA 5506
 Qy 4804 TTATATCATTAACGGAACTAAAGCGATTAATTCGCAAAAGTGTACTAAATTTGACCA 4863
 Db 5507 TTATATCATTAACGGAACTAAAGCGATTAATTCGCAAAAGTGTACTAAATTTGACCA 5566
 Qy 4864 ACTCGCAATCTCTCATCTTGCATCAGTGTGAAGCTTTTAAACAAGCTCTTAACATTA 4923
 Db 5567 ACTCGCAATCTCTCATCTTGCATCAGTGTGAAGCTTTTAAACAAGCTCTTAACATTA 5626
 Qy 4924 ACGCTAATCTGTTTAATTTGAAGCTATTAACGCTCTTTATGGAACCTGAGAGTGT 4983
 Db 5627 ACGCTAATCTGTTTAATTTGAAGCTATTAACGCTCTTTATGGAACCTGAGAGTGT 5686
 Qy 5044 TCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5103
 Db 5747 TCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5806
 Qy 5104 TAAATACAGAGCCCAACAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 5145
 Db 5807 TAAATACAGAGCCCAACAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 5848

RESULT 5
 AY069579
 LOCUS 5962 bp mRNA linear INV 17-DEC-2001
 DEFINITION Drosophila melanogaster LD31673 full length cDNA.
 ACCESSION AY069579
 VERSION AY069579.1 GI:17862493
 KEYWORDS FLI, CDNA.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

REFERENCE 1 (bases 1 to 5962)
 AUTHORS Stephenson, M., Brokstein, P., Hong, L., Aspayani, A., Carlson, J.,
 Champ, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R.,
 Gonzalez, M., Guarin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J.,
 Nuno, J., Pated, J., Paragas, V., Park, S., Phouanavong, S., Wan, K.,
 Yu, C., Lewis, S. E., Rubin, G. M. and Celisner, S.
 TITLE Direct Submission
 JOURNAL Submitted (10-DEC-2001) Berkeley Drosophila Genome Project,
 Lawrence Berkeley National Laboratory, One Cyclotron Road,
 Berkeley, CA 94720, USA
 COMMENT Sequence submitted by:
 Lawrence Berkeley National Laboratory
 Berkeley, CA 94720
 This clone was sequenced as part of a high-throughput process to
 sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
 Science 2000). The sequence has been subjected to integrity checks
 for sequence accuracy, presence of a polyA tail and contiguity
 within 100 kb in the genome. Thus we believe the sequence to
 reflect accurately this particular cDNA clone. However, there are
 artifacts associated with the generation of cDNA clones that may
 have not been detected in our initial analyses such as internal
 priming, priming from contaminating genomic DNA, retained introns
 due to reverse transcriptase errors that result in single base changes,
 and reverse transcriptase errors that result in single base changes.
 For further information about this sequence, including its location
 and relationship to other sequences, please visit our Web site
 (http://fruitfly.berkeley.edu) or send email to
 chma@fruitfly.berkeley.edu.
 FEATURES
 source Location/Qualifiers
 1. 5962
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /strain="y; cn bw sp"
 /db_xref="taxon:7227"
 1. 5961
 /note="blastn alignment with AE003593"
 766..5241
 /note="longest ORF"
 /codon_start=1
 /product="Ud31673p"
 /protein_id="PAL39724.1"
 /db_xref="GI:17862494"
 /translation="MAYRKPESDLDGPIQOMPKADMRVQVLAEDLVTFPSDTSYVC
 TDWGLIDGLMPWLQSHFKYAKSLKSLFSLKSLDPAATVATLPHVIRLGS
 RDTYREKALQILDMHRVLPQALIDKATSCFKKNAKREFLOTTVNALHETG
 TQOLSVVYIPVICALGDPVNVREAIOQLVEIKHVGRLPDLRMDVPSKX
 AMLEOKDOVKQEGILLPSALKNTNGVGLDEADNIGLBERPTMKRPLHSVSS
 LRPKNVNDVTGDAVATMESFESEFVPLNIFHAKMDMDIKOVLIVISDNAM
 EKRYDALKKIALLILSYHOPFOFVAOLKELSPDLIKELERLROVIEACTIY
 MSKTRKRLDAFCMSILEHINLONAKVIASTADLIKKIYKTAAPLKIYTT
 LNSKSDKIRSTLCLEMLVLFEEWOTAKENATVLDLTKSLGDDCAKRRSRTA
 YWARRRHPPELADISIGTLIDLAQALREEREGGGGTGTGTATETRTVRIK
 TPQTLQKPTPSMRISAVDTIAQRAKRAYTLYSRORKEPLGNNNSQASMTCAAS
 GSLRPLNNSGCTPATTPGSAVTPRGRAGVSOQSGASTSPYKLDQGGIAN
 YRGATGAIPKKAAGIPRSTASRSTSPRSGGLMKRSMYTAGSRTPERNNPPYR
 PSAARLLAAGREHNLGYCDGDPDYPVSGDWMRSGGMGRGLKLRDSDSDSA
 SSVCSESFDSSYTRKNSYSLSGSHTRLDMSRTORAPFDITITTOPCASTHSEK
 DGLISLQVYADKELTQOOLCVLDMFRKPMDFTHKRVSLFDVTELILVHANL
 HEWLFILTRILFNLGTDLLNSHMSIKMLQVVEYFPQOLKELFRILISSTORP
 TTKTRIALIRFLDNLANTYCKSDPESDOSACERTVLAQLAADKSMELSOAS
 CLVALNLTPOMTLLADLPKYODSARCSHSHMRQOSCSGANSPPSSLS
 SPKLOSPVGPASLOSHHOLISSTSPRSOSVSEOLLEPSELIDIONLOKTE
 EIRHCGQYOTALPANGNHLQYHOGQODSCASLSNKSOTSSANTOSNTPESA
 TMRLDNLERETONAKSPIDDACVITVSIIMANGELILASINMESEVVALTLTK
 DQVELLOTSLTNIGICIGKNCCLPNKGFPSIMRLMLNLEAHTVLAGLVSK
 IMRKNRMMHMFLELILKIKIOYHSHKALRDIDSMIPRIAPSLDISINIVP
 VIATGEPNNICAKILEVEHEHSEITDHLDIVPNNLARSADTOSMRKAVPC
 IVKLYFLGEEKVKKXSVLNPSTVRLNLYIEKORNCISGGSSTNSAASS"

ORIGIN
Query Match

98.3%; Score 5055.6; DB 2; Length 5962;

	Best Local Similarity	99.3%;	Pred. No. 0;	Matches 5108;	Conservative 0;	Mismatches 29;	Indels 5;	Gaps 3;
Qy	6	CGCAGCAGCGCCGCTCCGAATTTGTCGCCAGAACGAGAGCAATAGACATTGCAACG	65					
Db	706	CACGCGACGCCGATTTGCAATATGTTCCGCCAGAACGAGCAATAGACATTGCAACG	765					
Qy	66	ATGCGCTATCGAAGCCACGACCTTGATGCTTTATCCAGCAATGCCAAGCGGAC	125					
Db	766	ATGCGCTATCGAAGCCACGACCTTGATGCTTTATCCAGCAATGCCAAGCGGAC	825					
Qy	126	ATGCGCTATCGAAGCCACGACCTTGATGCTTTATCCAGCAATGCCAAGCGGAC	185					
Db	826	ATGCGCTATCGAAGCCACGACCTTGATGCTTTATCCAGCAATGCCAAGCGGAC	885					
Qy	186	TCAATTGCTCGACGAGCATGGGATTCCTTATGACGCTTATGATGCAATGCTGACGGC	245					
Db	886	TCAATTGCTCGACGAGCATGGGATTCCTTATGACGCTTATGATGCAATGCTGACGGC	945					
Qy	246	AGCCACTTTAAGATTGCAAAAGTCCCTGAGAGCGCTTCGGAGCTAATPAAGCATTTG	305					
Db	946	AGCCACTTTAAGATTGCAAAAGTCCCTGAGAGCGCTTCGGAGCTAATPAAGCATTTG	1005					
Qy	306	GCGACGATTTTATGATGATACAGGCTACCGCTTCGCAATGATGCAATCGGCTGGA	365					
Db	1006	GCGACGATTTTATGATGATACAGGCTACCGCTTCGCAATGATGCAATCGGCTGGA	1065					
Qy	366	GACAGCAGGACACAGTCCGAGAGGCGCACTTCTGCGCGCACTCATAGGACAC	425					
Db	1066	GACAGCAGGACACAGTCCGAGAGGCGCACTTCTGCGCGCACTCATAGGACAC	1125					
Qy	426	AGAGTCTTCGCGCCAGAGGCGCTGATGACAGAGCTGACCTGCTTCAAGCACAG	485					
Db	1126	AGAGTCTTCGCGCCAGAGGCGCTGATGACAGAGCTGACCTGCTTCAAGCACAG	1185					
Qy	486	AAGGCCAAGTGGCCGAGAGGATTCCTTACAGACATTTGTAACCTCTTACATGATGAGC	545					
Db	1186	AAGGCCAAGTGGCCGAGAGGATTCCTTACAGACATTTGTAACCTCTTACATGATGAGC	1245					
Qy	546	ACCAGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	605					
Db	1246	ACCAGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1305					
Qy	606	ACAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	665					
Db	1306	ACAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1365					
Qy	666	GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	725					
Db	1366	GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1425					
Qy	726	TGAGAGAAAGTTCAGCAGGCTCAACAGAGGAGGCTACTGCTACCTGAGCCCTTAA	785					
Db	1426	TGAGAGAAAGTTCAGCAGGCTCAACAGAGGAGGCTACTGCTACCTGAGCCCTTAA	1485					
Qy	786	AACAGCAATGCGCATGAGAGGCTTGAACAGAGCGCAATATTTGGTTGAGGAGCGA	845					
Db	1486	AACAGCAATGCGCATGAGAGGCTTGAACAGAGCGCAATATTTGGTTGAGGAGCGA	1545					
Qy	846	CCCAACAGATGATTTAAGCGGCGCATACATCGGCGGTTTGTGATCATCGGCCCAAAA	905					
Db	1546	CCCAACAGATGATTTAAGCGGCGCATACATCGGCGGTTTGTGATCATCGGCCCAAAA	1605					
Qy	906	CCCAATGTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	965					
Db	1606	CCCAATGTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1665					
Qy	966	AGCTTGAAGTGTGCTCCGCAATTTGAACATCTTTCACGCTAAGGACATGACATATCTAC	1025					
Db	1666	AGCTTGAAGTGTGCTCCGCAATTTGAACATCTTTCACGCTAAGGACATGACATATCTAC	1725					
Qy	1026	AAGCAAGTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1085					

Dh 1726 AAGCAAGTACTAGTATCATCAGTATATAAAACGACACTGGAGAAAAGTGTGATGCT 1785
Qy 1086 CTCAAGAGATCAGGGCAATGCTCATTCCTAGCTATCAACTCAGCCGAGTTTGCT 1145
Dh 1786 CTCAGAAAGATCAGGGCAATGCTCATTCCTAGCTATCAACTCAGCCGAGTTTGCT 1845
Qy 1146 GTACAGCTAAAGAAATGTCGTTAAGCTTGTGACATCTCTCAAGAGAACTACATCA 1205
Dh 1846 GTACAGCTAAAGAAATGTCGTTAAGCTTGTGACATCTCTCAAGAGAACTACATCA 1905
Qy 1206 CAGGTGATCCCGGAGAGCGTGCATCACTCGCTTCATGTCTAAGACCTGAGAAATAA 1265
Dh 1906 CAGGTGATCCCGGAGAGCGTGCATCACTCGCTTCATGTCTAAGACCTGAGAAATAA 1965
Qy 1266 CTAGATGCTTCTGTGTGAGACATTTTGGAGCACTGATTAATTAATCAGAACAGCCG 1325
Dh 1966 CTAGATGCTTCTGTGTGAGACATTTTGGAGCACTGATTAATTAATCAGAACAGCCG 2025
Qy 1326 AAGTCATTGATCGGCTTCCCAAAATAGCTGAAATATCATTAAGTATACATGCA 1385
Dh 2026 AAGTCATTGATCGGCTTCCCAAAATAGCTGAAATATCATTAAGTATACATGCA 2085
Qy 1386 CCAAAAGCTGCTCAAGATCTACACAGACACTCTGAATCAATCAAAAGTCCAAAGACATAAG 1445
Dh 2086 CCAAAAGCTGCTCAAGATCTACACAGACACTCTGAATCAATCAAAAGTCCAAAGACATAAG 2145
Qy 1446 TCCACACTGTGTGAGCTGATGGTGTGCTCTTCCAGAGAGTGGACAGAACAGCGTTGAA 1505
Dh 2146 TCCACACTGTGTGAGCTGATGGTGTGCTCTTCCAGAGAGTGGACAGAACAGCGTTGAA 2205
Qy 1506 AGGAATGCGACCGTACTAAGGAGACACTTAAATAATCCATTGGCGATGACAGCTGCAT 1565
Dh 2206 AGGAATGCGACCGTACTAAGGAGACACTTAAATAATCCATTGGCGATGACAGCTGCAT 2265
Qy 1566 GCAGCGCCGCTTCCAGATAGCGCTATGGGCTTTCAAGCGTCACTTTCCAGAGCTGGCG 1625
Dh 2266 GCAGCGCCGCTTCCAGATAGCGCTATGGGCTTTCAAGCGTCACTTTCCAGAGCTGGCG 2325
Qy 1626 GATCAATATATATGGAACATTAGACATAGCTGCGCCAGCGCATTTGAAAGGAAAGAG 1685
Dh 2326 GATCAATATATATGGAACATTAGACATAGCTGCGCCAGCGCATTTGAAAGGAAAGAG 2385
Qy 1686 GCGCGCGGAGAGAGAGAACTGTATCTGGGACTGGGACTGCACTGAAACGAGCGCT 1745
Dh 2386 GCGCGCGGAGAGAGAGAACTGTATCTGGGACTGGGACTGCACTGAAACGAGCGCT 2445
Qy 1746 GTATCCGCGATTTGGCCGAAACACCTGGAAACCTGCAAAAGCCGCTAATGATGATCC 1805
Dh 2446 GTATCCGCGATTTGGCCGAAACACCTGGAAACCTGCAAAAGCCGCTAATGATGATCC 2505
Qy 1806 ATTTCAAGCGGTGACACTGCGGCTGCTCAACGAGCGCAAGTTAGAGCGCAATATACATA 1865
Dh 2506 ATTTCAAGCGGTGACACTGCGGCTGCTCAACGAGCGCAAGTTAGAGCGCAATATACATA 2565
Qy 1866 TATTTCAAGCAAAAGAAACCTTTAGAGACTTAATAATTTCCAAACGAGCATCATGACAGGA 1925
Dh 2566 TATTTCAAGCAAAAGAAACCTTTAGAGACTTAATAATTTCCAAACGAGCATCATGACAGGA 2625
Qy 1926 GCAGACGCGATGATCACTGCGCCAGACCTGCGCTGAATTTCCATNAGGGGTGGACACCA 1985
Dh 2626 GCAGACGCGATGATCACTGCGCCAGACCTGCGCTGAATTTCCATNAGGGGTGGACACCA 2685
Qy 1986 GCTACAAAGCGCGGATCGGTTTACACACAGCGCCCGGGAGCGGACGAGATCTCAGTCA 2045
Dh 2686 GCTACAAAGCGCGGATCGGTTTACACACAGCGCCCGGGAGCGGACGAGATCTCAGTCA 2745
Qy 2046 CAGCCAGGATGAGATCCACTTCCCAAGCACAAAACTGAGGGATCAATGCGTGTGATTT 2105
Dh 2746 CAGCCAGGATGAGATCCACTTCCCAAGCACAAAACTGAGGGATCAATGCGTGTGATTT 2805
Qy 2106 GGTATTTACTATCTGTGGGGCCACTGGCGCCATATCCCAAAAAGGCTCCGGAAATACCCCGA 2165
Dh 2806 GGTATTTACTATCTGTGGGGCCACTGGCGCCATATCCCAAAAAGGCTCCGGAAATACCCCGA 2865

Qy 2166 AGCACAGCAGCTCTAGAGAAACGAGTCCAACTAGTCAAGTGTGGCTTGATGAAACGC 2225
Dh 2866 AGCACAGCAGCTCTAGAGAAACGAGTCCAACTAGTCAAGTGTGGCTTGATGAAACGC 2925
Qy 2226 AGTATGATCTTACAGGTGGGGGTCTTCACGATACGCCCGAGAGAAACAAACCGTAAAGA 2285
Dh 2926 AGTATGATCTTACAGGTGGGGGTCTTCACGATACGCCCGAGAGAAACAAACCGTAAAGA 2985
Qy 2286 CCATCGCGCGGACAGACTGTGGCGCAATCCGTAAGCAAGCACTAATTAAGCGCTC 2345
Dh 2986 CCATCGCGCGGACAGACTGTGGCGCAATCCGTAAGCAAGCACTAATTAAGCGCTC 3045
Qy 2346 GAGATGATGAGACAGCCAGACTATGTTTCCGGGAGCTAACGCGACGCGCGAATCGG 2405
Dh 3046 GAGATGATGAGACAGCCAGACTATGTTTCCGGGAGCTAACGCGACGCGCGAATCGG 3105
Qy 2406 ATGGGTAGGAAGCTCATGGGACGTGATGATGATGATGATGATGATGATGATGATGATGAT 2465
Dh 3106 ATGGGTAGGAAGCTCATGGGACGTGATGATGATGATGATGATGATGATGATGATGATGAT 3165
Qy 2466 GTGTGTCAGAACGATCTTTCAGACTCCAGCTACACTAGAGTAAATGAACTACTCA 2525
Dh 3166 GTGTGTCAGAACGATCTTTCAGACTCCAGCTACACTAGAGTAAATGAACTACTCA 3225
Qy 2526 CTTAGCGGAGCCACACCGGCTTGGACTGGAGACGAGCGGGCGCCATTGAGACATC 2585
Dh 3226 CTTAGCGGAGCCACACCGGCTTGGACTGGAGACGAGCGGGCGCCATTGAGACATC 3285
Qy 2586 GAGACGATTAATTCAGTTCTGCGCATGACGCAATGGCTGAAAGGATGCGCTGATC 2645
Dh 3286 GAGACGATTAATTCAGTTCTGCGCATGACGCAATGGCTGAAAGGATGCGCTGATC 3345
Qy 2646 AGCCTCACACGATCTGCGCCGATGGAAGAGACTCACTCAGAGCACTTAATGCGTT 2705
Dh 3346 AGCCTCACACGATCTGCGCCGATGGAAGAGACTCACTCAGAGCACTTAATGCGTT 3405
Qy 2706 CTGACACTGTTCCGCAAGATGTTCACTGATATACACACCAAGGATGATCACTGTTTTTG 2765
Dh 3406 CTGACACTGTTCCGCAAGATGTTCACTGATATACACACCAAGGATGATCACTGTTTTTG 3465
Qy 2766 GACACGCTCACGAACTGATTCGTGTTCAATGCAAGCAAACTCAGAAATGCGCTTCA 2825
Dh 3466 GACACGCTCACGAACTGATTCGTGTTCAATGCAAGCAAACTCAGAAATGCGCTTCA 3525
Qy 2826 TCTTGTGAGCGGCTTCAACAAATTTGGGACCGATCTACTCAATTCGATCAGACAGC 2885
Dh 3524 TCTTGTGAGCGGCTTCAACAAATTTGGGACCGATCTACTCAATTCGATCAGACAGC 3582
Qy 2886 AAGATTTGGAAGACGTTACAGTGTGTCACGAAATTTTCCACGCGAGTTGACCTTAA 2945
Dh 3583 AAGATTTGGAAGACGTTACAGTGTGTCACGAAATTTTCCACGCGAGTTGACCTTAA 3642
Qy 2946 GAACTATTTGATCAATATGATTTCTACTCAACGCACTCAAGACGCGGATGCGC 3005
Dh 3643 GAACTATTTGATCAATATGATTTCTACTCAACGCACTCAAGACGCGGATGCGC 3702
Qy 3006 ATTTGCGGCTTCTTACAGATCTGGCCATATACATCTGTAAGAGCAGCACTTTCCAGC 3065
Dh 3703 ATTTGCGGCTTCTTACAGATCTGGCCATATACATCTGTAAGAGCAGCACTTTCCAGC 3762
Qy 3066 GACCAGACGAGCGCTGCGAGCGGACGCTCTAAAGCTGCGCCAGCTGGCGCGGATGAG 3125
Dh 3763 GACCAGACGAGCGCTGCGAGCGGACGCTCTAAAGCTGCGCCAGCTGGCGCGGATGAG 3822
Qy 3126 AAGTCATGAGAGTGGCTGCGCCAGGACGAGAGCTGCTAGTGGCCCTGTAATACCTGAAT 3185
Dh 3823 AAGTCATGAGAGTGGCTGCGCCAGGACGAGAGCTGCTAGTGGCCCTGTAATACCTGAAT 3882
Qy 3186 ACCCGCAATGACCCCTTTACTGCGCAGACTGCAAGAGTATATCAGACTTGGCCGA 3245
Dh 3883 ACCCGCAATGACCCCTTTACTGCGCAGACTGCAAGAGTATATCAGACTTGGCCGA 3942

QY 3246 TCTGCATCATGCGCATGAGGCGGAAAGCCAAAGTTGCAATTCGGGTGCCAATTCG 3305
 |||||
 Db 3943 TCTGCATCATTCACACATGAGGCGGCAAAAGTTGCAATTCGGGTGCCAATTCG 4002
 QY 3306 CCTAGTACTCTCCATTGAGCAGTAGCAGTCCCAAGCTTTGCAAAAGTCCCTCTGGGT 3365
 |||||
 Db 4003 CCTAGTACTCTCCATTGAGCAGTAGCAGTCCCAAGCTTTGCAAAAGTCCCTCTGGGT 4062
 QY 3366 CCATTTGCTCGCTTCAGAGCCACCACTTAAAGATCAGCTCTTAAATGTCACGC 3425
 |||||
 Db 4063 CCATTTGCTCGCTTCAGAGCCACCACTTAAAGATCAGCTCTTAAATGTCACGC 4122
 QY 3426 TCCCGGAGCTCTTCGCTGAGAGAGAGCTCTTTCTCGGAGCTGAGCATTTAGAC 3485
 |||||
 Db 4123 TCCCGGAGCTCTTCGCTGAGAGAGAGCTCTTTCTCGGAGCTGAGCATTTAGAC 4182
 QY 3486 AACATTCAGAAAGCTCGGAGAGAGATCCGGGCACTGCTTCGGGCGTCAGTACAGAGCGG 3545
 |||||
 Db 4183 AACATTCAGAAAGCTCGGAGAGAGATCCGGGCACTGCTTCGGGCGTCAGTACAGAGCGG 4242
 QY 3546 CTGGCGCCCAATGGCTTCATGACATTCGATGATACAGATCAGGCGCAACAGGATTCG 3605
 |||||
 Db 4243 CTGGCGCCCAATGGCTTCATGACATTCGATGATACAGATCAGGCGCAACAGGATTCG 4302
 QY 3606 TGTGCAATCCCTGTCTTCCAACTCCAAAGCGCATCTGTGGCCAACTAACCCAGTCAAT 3665
 |||||
 Db 4303 TGTGCAATCCCTGTCTTCCAACTCCAAAGCGCATCTGTGGCCAACTAACCCAGTCAAT 4362
 QY 3666 ACACCTGATGAGCAACATGAGGCTGATATATCTGAGCGGGAAGGAAAGCACTCAGAC 3725
 |||||
 Db 4363 ACACCTGATGAGCAACATGAGGCTGATATATCTGAGCGGGAAGGAAAGCACTCAGAC 4422
 QY 3726 GCCAAGTCAACCACTGACGATGCGCAAGTGCATCGCTCGATTAATATGCTGAAAT 3785
 |||||
 Db 4423 GCCAAGTCAACCACTGACGATGCGCAAGTGCATCGCTCGATTAATATGCTGAAAT 4482
 QY 3786 GGAAGCTGATATCTGGCCAGACACTGATGAGAGCGAAGTGTGCTGTGGCTTACG 3845
 |||||
 Db 4483 GGAAGCTGATATCTGGCCAGACACTGATGAGAGCGAAGTGTGCTGTGGCTTACG 4542
 QY 3846 CTAACAAAGATTCAGCCCGCTGAGTGTGCTTACAGAGTCACTTAACCTGGGATTTGC 3905
 |||||
 Db 4543 CTAACAAAGATTCAGCCCGCTGAGTGTGCTTACAGAGTCACTTAACCTGGGATTTGC 4602
 QY 3906 ATCAAGGCGGAACTGTGAGCTGCCAATAGCACTTAATGATGATGCGAATGCTG 3965
 |||||
 Db 4603 ATCAAGGCGGAACTGTGAGCTGCCAATAGCACTTAATGATGATGCGAATGCTG 4662
 QY 3966 CTTAATCTTGAAGGCGGAGCATACGAGCGTGTGATCGCTGGCCTTGACGTCCTCAGT 4025
 |||||
 Db 4663 CTTAATCTTGAAGGCGGAGCATACGAGCGTGTGATCGCTGGCCTTGACGTCCTCAGT 4722
 QY 4026 AAGATTATGAGGAGCAAAATGCTGACAACTGAGTGCATTTCTAGAGCTGATTTG 4085
 |||||
 Db 4723 AAGATTATGAGGAGCAAAATGCTGACAACTGAGTGCATTTCTAGAGCTGATTTG 4782
 QY 4086 CTGAAGATCATCCAGTGTATCAACACAGCAAGAGGCTTTGGCGGATATGATGCTCATG 4145
 |||||
 Db 4783 CTGAAGATCATCCAGTGTATCAACACAGCAAGAGGCTTTGGCGGATATGATGCTCATG 4842
 QY 4146 ATACCAAGATAGACCATCTTGTGCTGTGATCTGTCATCAACATTTGATCCGTA 4205
 |||||
 Db 4843 ATACCAAGATAGACCATCTTGTGCTGTGATCTGTCATCAACATTTGATCCGTA 4902
 QY 4206 ATTGCGAGCGGTGATTTCCCAAAATTTGCGCCCATCAAGATCTGCTGAGGTGAC 4265
 |||||
 Db 4903 ATTGCGAGCGGTGATTTCCCAAAATTTGCGCCCATCAAGATCTGCTGAGGTGAC 4962
 QY 4266 GAACACATGTTGGAGATCACAGAGCGCCACTTGAGCATTTGTTCCCAATCTGGCG 4325
 |||||
 Db 4963 GAACACATGTTGGAGATCACAGAGCGCCACTTGAGCATTTGTTCCCAATCTGGCG 5022
 QY 4326 CGATGGCGGACGACAGCAATGATGTTGCAAAAGCTGGGCTTCTGATGATGTCAG 4385

Db 5023 CGATGGGAGACGACAGCAATGATGTTGCAAAAGCGGCTTCTGATGCTTAAG 5082
 |||||
 QY 4386 CTGACTTTTGTCTGGCGAGAGAGATCAAGCGGAGTGTGATGCTTAATCCAGC 4445
 |||||
 Db 5083 CTGACTTTTGTCTGGCGAGAGAGATCAAGCGGAGTGTGATGCTTAATCCAGC 5142
 QY 4446 AAGTTAGGCTCTCAACGTGTACATGAGAGAGCGGAACTGATCAGTGGGAGAG 4505
 |||||
 Db 5143 AAGTTAGGCTCTCAACGTGTACATGAGAGAGCGGAACTGATCAGTGGGAGAG 5202
 QY 4506 AGCTCTCAAAAGACCTCTCGCGGAGTGTGATGATTTGGGGAGCCCTTAATAGGA 4565
 |||||
 Db 5203 AGCTCTCAAAAGACCTCTCGCGGAGTGTGATGATTTGGGGAGCCCTTAATAGGA 5262
 QY 4566 TTCCTGCG--TCGTGCGACCAAAACAAGACACAGACCGGTTGCTCCCTGGCGTGA 4623
 |||||
 Db 5263 TTCCTGATGACAGACACACAAACAGACAGACAGACCGGTTGCTCCCTGGCGTGA 5322
 QY 4624 AGAAGTGAAGAGCGCGGACATTAATTAATTAATTAATTAATTAATTAATTA 4683
 |||||
 Db 5323 AGAAGTGAAGAGCGCGGACATTAATTAATTAATTAATTAATTAATTAATTA 5382
 QY 4684 CTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4743
 |||||
 Db 5383 CTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5442
 QY 4744 CAATGGGCGCGGAGAGAGAGTCAAGGCGGATGCGCCATCAATGATGAGAGAACTA 4803
 |||||
 Db 5443 CAATGGGCGCGGAGAGAGAGTCAAGGCGGATGCGCCATCAATGATGAGAGAACTA 5502
 QY 4804 TTATATCATTAACGCGGAACTAAAGCGAGTAATTCGCAAGGTGATTAATTTAGCA 4863
 |||||
 Db 5503 TTATATCATTAACGCGGAACTAAAGCGAGTAATTCGCAAGGTGATTAATTTAGCA 5562
 QY 4864 ACTCCGCAATCTCACTTCTCACTCATCATGTTGATGATGTTTATACAGCTCTTAATTA 4923
 |||||
 Db 5563 ACTCCGCAATCTCACTTCTCACTCATCATGTTGATGATGTTTATACAGCTCTTAATTA 5622
 QY 4924 ACGCTGATCTGTTAATTTGTAAGCTATATACCGCTCTTATGAGAACTCAGACGTGT 4983
 |||||
 Db 5623 ACGCTGATCTGTTAATTTGTAAGCTATATACCGCTCTTATGAGAACTCAGACGTGT 5682
 QY 4984 GCTTTTACCTTTGTTTCAAGGAGCTTTGATCTTATTTGTTTCCGCTCAACTAAT 5043
 |||||
 Db 5683 GCTTTTACCTTTGTTTCAAGGAGCTTTGATCTTATTTGTTTCCGCTCAACTAAT 5742
 QY 5044 TCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5103
 |||||
 Db 5743 TCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5802
 QY 5104 TAAATACAAAGCCACAAAGGATTAATTAATTAATTAATTAATTAATTAATTA 5145
 |||||
 Db 5803 TAAATACAAAGCCACAAAGGATTAATTAATTAATTAATTAATTAATTAATTA 5844

RESULT 6
 AC014071/c 87197 bp DNA linear HTG 16-NOV-1999
 LOCUS Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
 DEFINITION AC014071
 ACCESSION AC014071.1 GI:6437264
 VERSION HTG: HTGS PHASE2
 KEYWORDS Drosophila melanogaster (fruit fly)
 SOURCE Drosophila melanogaster
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 87197)
 AUTHORS Adams, M. and Venter, J.C.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
 Rockville, MD, USA

COMMENT

This sequence was identified as CDM:10214235 by the submitter.
For further information on this sequence e-mail to fly@celera.com.

- * NOTE: This is a 'working draft' sequence.
- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.

FEATURES

source
1. 87197
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="exon:7227"

ORIGIN

Query Match 98.2%; Score 5051.8; DB 14; Length 87197;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 5097; Conservative 0; Mismatches 17; Indels 5; Gaps 3;

```
OY 29 GTTGGCCGAGAACGAGACGATGACATTGCAACGATGCGCTTTCGAGAGCCGACGCA 88
Db 65836 GTTGGCCGAGAACGAGACGATGACATTGCAACGATGCGCTTTCGAGAGCCGACGCA
OY 89 CCTGATGAGCTTATCCAGCAAAATGCCAAGGCGGACATGCGTGTGAAGGTACAGCTCGC 148
Db 65776 CCTGATGAGCTTATCCAGCAAAATGCCAAGGCGGACATGCGTGTGAAGGTACAGCTCGC
OY 149 CGAGATCTGTGATCTCTTAGCGACACAAACTCAATTTGTCTGACCGGACATGGG 208
Db 65716 CGAGATCTGTGATCTCTTAGCGACACAAACTCAATTTGTCTGACCGGACATGGG
OY 209 ATTCTTTATGACGGTTTGATGCGCATGGCTGACGGGCGACCTTTAAGATTGCAAAA 268
Db 65656 ATTCTTTATGACGGTTTGATGCGCATGGCTGACGGGCGACCTTTAAGATTGCAAAA
OY 269 GTCCCTGAGAGGGCTCTCGAGACTAATAAGCGATTGGGCGACGATTTTAATGATACAC 328
Db 65536 GTCCCTGAGAGGGCTCTCGAGACTAATAAGCGATTGGGCGACGATTTTAATGATACAC
OY 329 GCGTACCGTTCTGCACATGTGATCGCTGGGAGACAGCAGGAGCACAGTCCGCGA 388
Db 65536 GCGTACCGTTCTGCACATGTGATCGCTGGGAGACAGCAGGAGCACAGTCCGCGA
OY 389 GAAGCGCAACTTCTGCTGCGGACCTCTAGAGACACAGAGTGTCTCCGCCAGCGCT 448
Db 65476 GAAGCGCAACTTCTGCTGCGGACCTCTAGAGACACAGAGTGTCTCCGCCAGCGCT
OY 449 GATCGACAAGCTGCGACATGCTGCTTCAAGCACAAGAACCGCAGAGTGGCGAGAGTT 508
Db 65416 GATCGACAAGCTGCGACATGCTGCTTCAAGCACAAGAACCGCAGAGTGGCGAGAGTT
OY 509 CTTTCAGACGATTTGTAACGCTCTTCATGAGTACGCGACCCAGCAGCTTGTTCGGGT 568
Db 65356 CTTTCAGACGATTTGTAACGCTCTTCATGAGTACGCGACCCAGCAGCTTGTTCGGGT
OY 569 CTATATACCAACGATTTGTGCACTTCTCGAGATCTCCACAGTTAATGTGAGGAGAGCGGC 628
Db 65236 CTATATACCAACGATTTGTGCACTTCTCGAGATCTCCACAGTTAATGTGAGGAGAGCGGC
OY 629 CATCCAAACGATGAGAAATCTACAGCATGTAGGGGATCGATTGGCGCCGACGCTCGC 688
Db 65236 CATCCAAACGATGAGAAATCTACAGCATGTAGGGGATCGATTGGCGCCGACGCTCGC
OY 689 TCGCATGAGCAGATGTTCTCGCTCGAAATTTGGCTATGTTGAGCAAAAGTTGACAGAGT 748
Db 65176 TCGCATGAGCAGATGTTCTCGCTCGAAATTTGGCTATGTTGAGCAAAAGTTGACAGAGT
OY 749 CAAACAGAGGCTCTACTGCTACTCTTCAGCCCTTAAAAACAGAAATGGAATGAGAGTGG 808
Db 65116 CAAACAGAGGCTCTACTGCTACTCTTCAGCCCTTAAAAACAGAAATGGAATGAGAGTGG
OY 809 CTTGACGAGGCGCAGCAATATTTGGTTGAGGAGGAGCCACCAAGATGATTTAAAGCGGC 868
Db 65056 CTTGACGAGGCGCAGCAATATTTGGTTGAGGAGGAGCCACCAAGATGATTTAAAGCGGC
OY 65056 CTTGACGAGGCGCAGCAATATTTGGTTGAGGAGGAGCCACCAAGATGATTTAAAGCGGC 64997
```

```
OY 869 ACTACATCTGGCCCTTTCTGTCATCTACCTGCGCCCAAAACCCAAATGTGAACGATGACCGG 928
Db 64996 ACTACATCTGGCCCTTTCTGTCATCTACCTGCGCCCAAAACCCAAATGTGAACGATGACCGG
OY 929 TGAATGCGGCGCGGTAAACGATGAAATCTTTGAAATCTAGCTTTGAGGTGTCGCCAATT 988
Db 64936 TGAATGCGGCGCGGTAAACGATGAAATCTTTGAAATCTAGCTTTGAGGTGTCGCCAATT
OY 989 GAACATCTTCCACGCTAAGGACATGACGATATCTACAAAGCACTAGTATGATCATCAG 1048
Db 64876 GAACATCTTCCACGCTAAGGACATGACGATATCTACAAAGCACTAGTATGATCATCAG
OY 1049 TGATAAAAAGCGACACTGGGAGAAAGCTGTGATGCTCTCAAGAAATCAGGGAATTCCT 1108
Db 64816 TGATAAAAAGCGACACTGGGAGAAAGCTGTGATGCTCTCAAGAAATCAGGGAATTCCT
OY 1109 CATTTGAGCTATGACACTGACGCGGAGTTTGTGCTGTACAGCTTAAGAAATTTGTCGT 1168
Db 64756 CATTTGAGCTATGACACTGACGCGGAGTTTGTGCTGTACAGCTTAAGAAATTTGTCGT
OY 1169 AAGCTTGTGACATCTCTCAAGAGAACTACATCAAGGTATCCGCGAGCGTGCAT 1228
Db 64696 AAGCTTGTGACATCTCTCAAGAGAACTACATCAAGGTATCCGCGAGCGTGCAT
OY 1229 CACCATGCTCTACATGCTCTAAGAGCGCTGAGAAATTAATCTAGTATGCTTGTGAGCAT 1288
Db 64636 CACCATGCTCTACATGCTCTAAGAGCGCTGAGAAATTAATCTAGTATGCTTGTGAGCAT
OY 1289 TTTGAGACACTGATTAATTAATTAATACAGAAACAGCGCGAAGGTCAATGATCCGCTTCAC 1348
Db 64576 TTTGAGACACTGATTAATTAATTAATTAATACAGAAACAGCGCGAAGGTCAATGATCCGCTTCAC
OY 1349 AATAGCTCTGAAGTATATCATTAAGTATACATGACACCAAGCTGCTCAAGATCTACAC 1408
Db 64516 AATAGCTCTGAAGTATATCATTAAGTATACATGACACCAAGCTGCTCAAGATCTACAC
OY 1409 AGACACTGTGAATCAATCAAAAGTGAAGACATTAAGGTCCACATGCTGATGATGATGCT 1468
Db 64456 AGACACTGTGAATCAATCAAAAGTGAAGACATTAAGGTCCACATGCTGATGATGATGCT
OY 1469 GCTGCTCTTCAAGAGAGTGGACAGACGAAAGCGTTGGAAGAAATCCACCTGATTAAGGGA 1528
Db 64396 GCTGCTCTTCAAGAGAGTGGACAGACGAAAGCGTTGGAAGAAATCCACCTGATTAAGGGA
OY 1529 CACCTTAATAAATTCATTTGGCGATGACACTGCGATGACCGCCGATTCAGATACGC 1588
Db 64336 CACCTTAATAAATTCATTTGGCGATGACACTGCGATGACCGCCGATTCAGATACGC
OY 1589 CTATTTGGGCTTTCAAGGCTCACTTTCCAGAGCTGGCGGATCAATATATGGAACATTGA 1648
Db 64276 CTATTTGGGCTTTCAAGGCTCACTTTCCAGAGCTGGCGGATCAATATATGGAACATTGA
OY 1649 CATAGCTGCGCAGCGCATTAGAAAGGAAAGGAGCGCGCGAGAGGAGGAGAACTGG 1708
Db 64216 CATAGCTGCGCAGCGCATTAGAAAGGAAAGGAGCGCGCGAGAGGAGGAGAACTGG
OY 1709 TACTGGAGCTGGGACTGACCTGTAAGACGACGCTGATCCCGCATTTGCCGAAACACC 1768
Db 64156 TACTGGAGCTGGGACTGACCTGTAAGACGACGCTGATCCCGCATTTGCCGAAACACC
OY 1769 TGGAAACCTGGAAGAGCGCGCTAGTATGAGATTCATTTCAAGGGGAGCATGCGGGC 1828
Db 64096 TGGAAACCTGGAAGAGCGCGCTAGTATGAGATTCATTTCAAGGGGAGCATGCGGGC
OY 1829 TGTCTCAAGAGCGCAAGTTAGAGCGCAATATACATATATTCAGGCAAGAAACCTTT 1888
Db 64036 TGTCTCAAGAGCGCAAGTTAGAGCGCAATATATACATATATTCAGGCAAGAAACCTTT
OY 1889 AGACCTTAATATTTCCAAACGAGCATGATGACAGAGCAGACGACGATGATCATCTGCC 1948
Db 63976 AGACCTTAATATTTCCAAACGAGCATGATGACAGAGCAGACGACGATGATCATCTGCC
OY 1949 CAGACCTGCGCTGAATTTCCAAATAGCGGTGGACACACGACTACAAACGCGGGATGGTTAC 2008
```

Db	631916	CAGACCTCGCTGAATTCCATTAGCGGTGGCACAACGCTTACAAACCGCCGGATCGGTTTAC	63857
Oy	2009	ACCAAGGCCCCGGGGACGGGACAGAGTATCTCACTCAACAGCAAGATCGAGATCCACTTC	2068
Db	63856	ACCAAGGCCCCGGGGACGGGACAGAGTATCTCACTCAACAGCAAGATCGAGATCCACTTC	63797
Oy	2069	GCCAAAGCAAAACTGAGGAGTCACTACGGTGGATTGGTAAATTACTATCTGTGGGCGAC	2128
Db	63796	GCCAAAGCAAAACTGAGGAGTCACTACGGTGGATTGGTAAATTACTATCTGTGGGCGAC	63737
Oy	2129	TGGCGCCATACCCAAAAAGGCGCTCCGGATATCCCGGAAGACAGCAGGCTCTAGGGAAAC	2188
Db	63736	TGGCGCCATACCCAAAAAGGCGCTCCGGATATCCCGGAAGACAGCAGGCTCTAGGGAAAC	63677
Oy	2189	GAGTCCAACTAGTCAAGTGTGGCTGTGATGAAACGAGATGTACTTACAGGTGCGG	2248
Db	63676	GAGTCCAACTAGTCAAGTGTGGCTGTGATGAAACGAGATGTACTTACAGGTGCGG	63617
Oy	2249	GTCTTGACGTACGCCCCGAGAGAACAAACCCAGTAAACATCGGCGCGGACGACTGCT	2308
Db	63616	GTCTTGACGTACGCCCCGAGAGAACAAACCCAGTAAACATCGGCGCGGACGACTGCT	63557
Oy	2309	GGCGCAATCCCGGTAAAGAGAACATATATAGGCTGGAGATGATGACAGCCAAACTA	2368
Db	63556	GGCGCAATCCCGGTAAAGAGAACATATATAGGCTGGAGATGATGACAGCCAAACTA	63497
Oy	2369	TGTTTCCGGGACTACATAGCGCACCGGCGGAATCCGATGGTAGGAAGCTCATGGGACG	2428
Db	63496	TGTTTCCGGGACTACATAGCGCACCGGCGGAATCCGATGGTAGGAAGCTCATGGGACG	63437
Oy	2429	TGATGAGTCTGATGACATCGACTTCGAGGCCAGTCTGTGTGTTCAAAACGATCTTTCGA	2488
Db	63436	TGATGAGTCTGATGACATCGACTTCGAGGCCAGTCTGTGTGTTCAAAACGATCTTTCGA	63377
Oy	2489	CTCCAGCTACCTAGAGTAAATTAATCGAATCTACTCTTAAGGGGAGCCACACCCGCTT	2548
Db	63376	CTCCAGCTACCTAGAGTAAATTAATCGAATCTACTCTTAAGGGGAGCCACACCCGCTT	63317
Oy	2549	GGACTGAGAGCACGAGCGGGGCGCAATTGACAGCATCGAGACGATTAATTCAGTCTGCGC	2608
Db	63316	GGACTGAGAGCACGAGCGGGGCGCAATTGACAGCATCGAGACGATTAATTCAGTCTGCGC	63257
Oy	2609	ATCGACGCAATTGATCTGAAAGGAAGATGGCTGATCGAGCTCACACAGTATCTGGCGA	2668
Db	63256	ATCGACGCAATTGATCTGAAAGGAAGATGGCTGATCGAGCTCACACAGTATCTGGCGA	63197
Oy	2669	TGGAAGGAGTCACTCAGACAGCACTTAAATGCGTTCTGGACATGTTCCGCAAGTGT	2728
Db	63196	TGGAAGGAGTCACTCAGACAGCACTTAAATGCGTTCTGGACATGTTCCGCAAGTGT	63137
Oy	2729	CATGAGATACACACCAAGGTTACTACGTTTTTTGACAGCGGTACCGAATCTGATTT	2788
Db	63136	CATGAGATACACACCAAGGTTACTACGTTTTTTGACAGCGGTACCGAATCTGATTT	63077
Oy	2789	GGTTCATGCGAACCAACTTCAGCAAAATGCGCTTCACTTGTGTTACCGCGTTGTTCAA	2848
Db	63076	GGTTCATGCGAACCAACTTCAGCAAAATGCGCTTCACTTGTGTTACCGCGTTGTTCAA	63020
Oy	2849	CAAAATGGGACCGGATCTCAATTCGATGCAACGACAAAGTTGGAAGCGCTACAGGT	2908
Db	63019	CAAAATGGGACCGGATCTCAATTCGATGCAACGACAAAGTTGGAAGCGCTACAGGT	62960
Oy	2909	GGTCCAGGAATTTTCCCAAGGCTTGACGCTTAAAGAACTATTAGAAATCATATCGGA	2968
Db	62959	GGTCCAGGAATTTTCCCAAGGCTTGACGCTTAAAGAACTATTAGAAATCATATCGGA	62900
Oy	2969	TTTCTACTCAACGCAACTACCAAGACGCGCATCGCATCTTGCGCTTCTTACGGATCT	3028
Db	62899	TTTCTACTCAACGCAACTACCAAGACGCGCATCGCATCTTGCGCTTCTTACGGATCT	62840
Oy	3029	GGCCATACATACGTGTAAGACAGCGATTTCCAGGACAGAGCCAGGCTGCGAGCG	3088

Db	62839	GGCCAAATACATCTGCAAAAGACGACCTTTCCACAGACGAGCAGGCTGGCAGGG	62789
Qy	3089	GACGGTCTTAAAGCTGGCCAGCTGGGCGCGATCGAAAGTGGATGGAGTGGCTCCCA	3148
Db	62779	GACGGTCTTAAAGCTGGCCAGCTGGGCGGATCGAAAGTGGATGGAGTGGCTCCCA	62720
Qy	3149	GGCCAGAGAGCTGCTAGTGGCCCTGATATACCGAATACCCGCAATGACCCCTTTACT	3208
Db	62719	GGCCAGAGAGCTGCTAGTGGCCCTGATATACCGAATACCCGCAATGACCCCTTTACT	6266
Qy	3209	GGCCGACCTGCAAAAGGTATATCAGGACTCTGCCGATCCCTGGATCCATTTCGACATGAG	3268
Db	62659	GGCCGACCTGCAAAAGGTATATCAGGACTCTGCCGATCCCTGGATCCATTTCGACATGAG	62600
Qy	3269	GGGCAAAAGCCAAAGTTGCAATTGGGAGTCGAATTGGCTAGTAGCTTCTTCATTGACAG	3328
Db	62599	GGGCAAAAGCCAAAGTTGCAATTGGGAGTCGAATTGGCTAGTAGCTTCTTCATTGACAG	62540
Qy	3329	TAGAGATGCCAAGCCTTTGGAAAGTCCCTGTGGGATCCATTGGCTCCCTTCAGAGCA	3388
Db	62539	TAGAGATGCCAAGCCTTTGGAAAGTCCCTGTGGGATCCATTGGCTCCCTTCAGAGCA	62480
Qy	3389	CCACACCAACCTTAGCATCAGCTCTCTACTAGTCCACGCTCCCGCAGTCTTCCGTGGAGCA	3448
Db	62479	CCACACCAACCTTAGCATCAGCTCTCTACTAGTCCACGCTCCCGCAGTCTTCCGTGGAGCA	62420
Qy	3449	GGAGCTCTCTTTTCTCTCGAGCTGAGCAATTGACGACCAATTTCAGAAAGCTGGAGGA	3508
Db	62419	GGAGCTCTCTTTTCTCTCGAGCTGAGCAATTGACGACCAATTTCAGAAAGCTGGAGGA	62360
Qy	3509	GATCCGGCATCTGCTTGGCGGCTCAGTACCAAGAGCGCTGGGGCCCATGGCTTCAATGG	3568
Db	62359	GATCCGGCATCTGCTTGGCGGCTCAGTACCAAGAGCGCTGGGGCCCATGGCTTCAATGG	62300
Qy	3569	ACACTTCACATATCAGATCAGGAGGCCAAGAGTTGCTGGATCCCTGCTTCCCACTC	3628
Db	62299	ACACTTCACATATCAGATCAGGAGGCCAAGAGTTGCTGGATCCCTGCTTCCCACTC	62240
Qy	3629	CAAGACCAATTCGTGGCCACACTACCCAGTCAAAATACACTGACTCAGCAACATGAG	3688
Db	62239	CAAGACCAATTCGTGGCCACACTACCCAGTCAAAATACACTGACTCAGCAACATGAG	62180
Qy	3689	GCTGGATATATCTGAGCGGGAAGGACCACTCAGAACGCCAAGTCAACCACTGACATGC	3748
Db	62179	GCTGGATATATCTGAGCGGGAAGGACCACTCAGAACGCCAAGTCAACCACTGACATGC	62120
Qy	3749	CAAGGTGATCAAGGTCCGATTAATATATGCTGAAATATGAGAGCTGATATCTGGCCAGCA	3808
Db	62119	CAAGGTGATCAAGGTCCGATTAATATATGCTGAAATATGAGAGCTGATATCTGGCCAGCA	62060
Qy	3809	CTTGATGAGAGGGAAGTGGTGGTGGCTCTTGCAGCTTAACGAAGAGTCAAGCCGTGCA	3868
Db	62059	CTTGATGAGAGGGAAGTGGTGGTGGCTCTTGCAGCTTAACGAAGAGTCAAGCCGTGCA	62000
Qy	3869	GTTGCTTCAGAGCTCACTTAATTAACCTGGGAGTTTGATCAAGGCGGAACTGTGAGCT	3928
Db	61999	GTTGCTTCAGAGCTCACTTAATTAACCTGGGAGTTTGATCAAGGCGGAACTGTGAGCT	61940
Qy	3929	GCCCAATTAAGCACTTTAAGATCCGATCATATGGGAGTGGCTTAACATTCTGGAAGCGAGCA	3988
Db	61939	GCCCAATTAAGCACTTTAAGATCCGATCATATGGGAGTGGCTTAACATTCTGGAAGCGAGCA	61880
Qy	3989	TACGAGAGTGTATCATCGCTGGCGCTGCAACGTGCTCAGTAAGATTATGAGAGCAACAAAT	4048
Db	61879	TACGAGAGTGTATCATCGCTGGCGCTGCAACGTGCTCAGTAAGATTATGAGAGCAACAAAT	61820
Qy	4049	GCGTCACACATGATGACCTTTCTAGAGCTGATTTTGTGAAAGATCATCAAGTGTATCA	4108
Db	61819	GCGTCACACATGATGACCTTTCTAGAGCTGATTTTGTGAAAGATCATCAAGTGTATCA	61760
Qy	4109	ACACAGGAAGAGAGCTTTGGGGATATGCACTCGATGATATCCAGATATGACACATCTCT	4168
Db	61759	ACACAGGAAGAGAGCTTTGGGGATATGCACTCGATGATATCCAGATATGACACATCTCT	617000

Accession	Gene	Position	Strand	Length	GC Content	GC Skew	GC Bias	GC Bias2	GC Bias4	GC Bias8	GC Bias16	GC Bias32	GC Bias64	GC Bias128	GC Bias256	GC Bias512	GC Bias1024	GC Bias2048	GC Bias4096	GC Bias8192	GC Bias16384	GC Bias32768	GC Bias65536	GC Bias131072	GC Bias262144	GC Bias524288	GC Bias1048576	GC Bias2097152	GC Bias4194304	GC Bias8388608	GC Bias16777216	GC Bias33554432	GC Bias67108864	GC Bias134217728	GC Bias268435456	GC Bias536870912	GC Bias1073741824	GC Bias2147483648	GC Bias4294967296	GC Bias8589934592	GC Bias17179869184	GC Bias34359738368	GC Bias68719476736	GC Bias137438953472	GC Bias274877906944	GC Bias549755813888	GC Bias1099511627776	GC Bias2199023255552	GC Bias4398046511104	GC Bias8796093022208	GC Bias17592186044416	GC Bias35184372088832	GC Bias70368744177664	GC Bias140737488355328	GC Bias281474976710656	GC Bias562949953421312	GC Bias1125899906842624	GC Bias2251799813685248	GC Bias4503599627370496	GC Bias9007199254740992	GC Bias18014398509481984	GC Bias36028797018963968	GC Bias72057594037927936	GC Bias144115188075855872	GC Bias288230376151711744	GC Bias576460752303423488	GC Bias1152921504606846976	GC Bias2305843009213693952	GC Bias4611686018427387904	GC Bias9223372036844775808	GC Bias18446744073689551616	GC Bias36893488147379103232	GC Bias73786976294758206464	GC Bias147573952589516412928	GC Bias295147905179032825856	GC Bias590295810358065651712	GC Bias1180591620716131303424	GC Bias2361183241432262606848	GC Bias4722366482864525213696	GC Bias9444732965729050427392	GC Bias18889465931458100844784	GC Bias377789318629162016896	GC Bias755578637258324033792	GC Bias1511157274516648067584	GC Bias3022314549033296135168	GC Bias6044629098066592270336	GC Bias12089258196133184540672	GC Bias24178516392266369081344	GC Bias48357032784532738162688	GC Bias96714065569065476325376	GC Bias193428131138130952650752	GC Bias386856262276261905301504	GC Bias773712524552523810603008	GC Bias1547425049105047621206016	GC Bias30948500982100952424121312	GC Bias61897001964201904848242624	GC Bias12379400392840380896485248	GC Bias24758800785680761792970496	GC Bias4951760157136152359584192	GC Bias9903520314312304719168384	GC Bias19807040628624609438337664	GC Bias39614081257249218876675328	GC Bias79228162514498437753350656	GC Bias15845632502899687550670112	GC Bias31691265005799375101340224	GC Bias63382530011598750202680448	GC Bias126765060023197504405368896	GC Bias253530120046395008810737792	GC Bias5070602400927900176214757824	GC Bias101412048018558003524289544	GC Bias202824096037116007048569088	GC Bias405648192074232014093777776	GC Bias811296384148464028187555552	GC Bias162259276829692805635111104	GC Bias324518553659385611270222208	GC Bias64903710721187722224444416	GC Bias12980742144375444488888832	GC Bias25961484288750888977777664	GC Bias5192296857751677795555328	GC Bias1038459371543355591111056	GC Bias207691874308671118222222224	GC Bias415383748617342224444448	GC Bias83076749723568448888896	GC Bias166153499447136897777776	GC Bias3323069988942737955555328	GC Bias6646139977885475911111056	GC Bias132922799557710951822222224	GC Bias265845599115418104444448	GC Bias531691198230836208888896	GC Bias1063382396461672417777776	GC Bias21267647929233448355555328	GC Bias42535295858466896711111056	GC Bias85070591716933793422222224	GC Bias1701411834338675684444448	GC Bias3402823668677351368888896	GC Bias680564733735273273777776	GC Bias136112946747054444755555328	GC Bias272225893494108889511111056	GC Bias54445178698821777822222224	GC Bias1088903739763355564444448	GC Bias21778074795267111288888896	GC Bias4355614959053422248888896	GC Bias871122991810684449777776	GC Bias17422498362013688955555328	GC Bias34844996724027377911111056	GC Bias69689993448054755822222224	GC Bias13937998689610911150444448	GC Bias27875997379221822308888896	GC Bias5575199475844364461777776	GC Bias11150398951688728128355555328	GC Bias2230079790337745625711111056	GC Bias4460159580675491251422222224	GC Bias892031916135188250284444448	GC Bias178406384270396450568888896	GC Bias35681276854079290113777776	GC Bias7136255370807858022755555328	GC Bias14272510741615716450511111056	GC Bias285450214
-----------	------	----------	--------	--------	------------	---------	---------	----------	----------	----------	-----------	-----------	-----------	------------	------------	------------	-------------	-------------	-------------	-------------	--------------	--------------	--------------	---------------	---------------	---------------	----------------	----------------	----------------	----------------	-----------------	-----------------	-----------------	------------------	------------------	------------------	-------------------	-------------------	-------------------	-------------------	--------------------	--------------------	--------------------	---------------------	---------------------	---------------------	----------------------	----------------------	----------------------	----------------------	-----------------------	-----------------------	-----------------------	------------------------	------------------------	------------------------	-------------------------	-------------------------	-------------------------	-------------------------	--------------------------	--------------------------	--------------------------	---------------------------	---------------------------	---------------------------	----------------------------	----------------------------	----------------------------	----------------------------	-----------------------------	-----------------------------	-----------------------------	------------------------------	------------------------------	------------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	--------------------------------	------------------------------	------------------------------	-------------------------------	-------------------------------	-------------------------------	--------------------------------	--------------------------------	--------------------------------	--------------------------------	---------------------------------	---------------------------------	---------------------------------	----------------------------------	-----------------------------------	-----------------------------------	-----------------------------------	-----------------------------------	----------------------------------	----------------------------------	-----------------------------------	-----------------------------------	-----------------------------------	-----------------------------------	-----------------------------------	-----------------------------------	------------------------------------	------------------------------------	-------------------------------------	------------------------------------	------------------------------------	------------------------------------	------------------------------------	------------------------------------	------------------------------------	-----------------------------------	-----------------------------------	-----------------------------------	----------------------------------	----------------------------------	------------------------------------	---------------------------------	--------------------------------	---------------------------------	----------------------------------	----------------------------------	------------------------------------	---------------------------------	---------------------------------	----------------------------------	-----------------------------------	-----------------------------------	-----------------------------------	----------------------------------	----------------------------------	---------------------------------	------------------------------------	------------------------------------	-----------------------------------	----------------------------------	-----------------------------------	----------------------------------	---------------------------------	-----------------------------------	-----------------------------------	-----------------------------------	-----------------------------------	-----------------------------------	----------------------------------	--------------------------------------	-------------------------------------	-------------------------------------	------------------------------------	------------------------------------	-----------------------------------	-------------------------------------	--------------------------------------	------------------

DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
<i>Drosophila melanogaster</i> 3L BAC RP98-24f3/Roswell Park Cancer Institute <i>Drosophila</i> BAC library complete sequence.	AC010055						
	AC010055						
	AC010055.6	GI:21392454					
	HTG.						
<i>Drosophila melanogaster</i> (fruit fly)							
<i>Drosophila melanogaster</i>							
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Pnydroidea; Drosophilidae; Drosophila.							
	1 (bases 1 to 163098)						
	Mazyu,D., Scherer,S., Adams,M.D., Holt,R.A., Evans,C.A.,						
	Gucay-Rech,S.D., Tabor,P., Williamson,A.H., Homsl,F.H.,						
	Dugan-Roch,S.D., Sodergren,E.S., Hodgson,A.H., Chen,R.C.,						
	Ayala,M., Scott,G.S., Morley,K.W., Amaratunga,P.G., Brandon,R.C.,						
	Rogers,T., An,H., Baldwin,D., Beeson,K.Y., Brown,M., Buhay,C.,						
	Busan,D.A., Canter,A., Chen,G., Chen,Z., Cleric-Blankenburg,K.,						
	Davenport,L.B., Dietz,S.M., Ding,Y., Dodson,K., Doup,L.E.,						
	Draper,H., Emery-Cohen,A., Ferreira,S., Gary,N.D.S., Houck,J.,						
	Hostin,D., Howland,T.J., Hume,J., Idegawa,C., Jajali,M., Kovar,C.,						
	Itu,M., Mattee,B., McIntosh,T.C., Morgan,M., Moy,M., Murphy,B.,						
	Nelson,K.A., Ndaasa,Y., Nguyen,N., Perez,L., Pitman,G.S., Puri,V.,						
	Schaefer,F., Shen,H., Strong,R., Tector,C., Wang,Q., Williams,S.M.,						
	Xiang,J., Zaveri,J.S., Zhou,J., Zorrilla,S., Smith,H.O.,						
	Wheeler,D., Weinstein,G., Gibbs,R. and Venter,J.C.						
	Direct Submission						
	Unpublished						
	2 (bases 1 to 163098)						
	Worley,K.C., Adams,C., Adio-Ogunola,B., Ali-osman,F.R., Allen,C.,						
	Alsbrooks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbarta,J.,						
	Benton,J., Blinage,K., Blankenburg,K., Bonini,D., Bouck,J.,						
	Bowie,S., Brieve,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,						
	Burton,P., Burrell,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,						
	Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,						
	Chen,Z., Chowhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,						
	Coyla,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,						
	Davy-Carroll,L., Dederich,H.A., Delaney,K.R., Delgado,O.,						
	Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,						
	Dugan-Rochs,S., Durbin,K.U., Earnhart,C., Edgar,D., Edwards,C.C.,						
	Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,						
	Forster,P., Frintz,P., Gabriel,A., Gao,J., Garcia,A., Garner,T.,						
	Garza,N., Gill,R., Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S.,						
	Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,						
	Hernandez,K., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,						
	Hollins,B., Homsl,F., Howard,S., Huber,J., Hulty,S., Hume,J.,						
	Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,						
	Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Koryah,J.,						
	Kovar,C., Kratovic,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C.,						
	Lewis,L., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,						
	Loulsgead,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,						
	Ma,J., Maheshwari,M., Mapua,P., Martin,R., Matindale,A.,						
	Marthner,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,						
	Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,						
	Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,						
	Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nnokenkwu,S.,						
	Ogun,M., Okunodu,G., Otagunye,N., Oyedero,R., Pace,A., Payton,B.,						
	Peery,J., Peters,L., Pickens,R., Primus,E., Pu,L.L.,						
	Quiles,M., Ren,Y., Rives,M., Rojals,A., Rojibabon,I., Royle,M.,						
	Ritz,S., Severy,G., Scherer,S., Scott,G., Shen,H., Shoochert,N.,						
	Sisson,I., Sodergren,E., Sonake,T., Sparks,A., Stanley,H.,						
	Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamaria,A., Tamaria,K.,						
	Tang,H., Tansey,U., Taylor,C., Taylor,T., Tellrod,B., Thomas,N.,						
	Thomas,S., Usumari,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,						
	Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,						
	Wellington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,						
	Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,						
	Weinstock,G. and Gibbs,R.						
	Direct Submission						
	Unpublished						
	3 (bases 1 to 163098)						
	Worley,K.C.						
	Direct Submission						
	Submitted (11-SEP-1999)						

Human Genome Sequencing Center, Department

QY 269 GTCCCTGAGGCGTTCTCGAGCTAATAAGCATTTGGCAGACGATTTTATGATACAC 328
Db 59511 GTCCCTGAGGCGTTCTCGAGCTAATAAGCATTTGGCAGACGATTTTATGATACAC 59512
QY 329 GGCTACCGTCTTGCCACATGATGATCGCTGGAGACAGCAGGACACAGTCCGGA 388
Db 59511 GGCTACCGTCTTGCCACATGATGATCGCTGGAGACAGCAGGACACAGTCCGGA 59472
QY 389 GAAGCGCAAGCTTCTGCTGCGGACCTCAATGAGACAGAGTCTCCGCCAGCGCT 448
Db 59471 GAAGCGCAAGCTTCTGCTGCGGACCTCAATGAGACAGAGTCTCCGCCAGCGCT 59412
QY 449 GATCAGCAAGCTGCGCAGCTAGCTGCTTCAAGCACAAGACCGCAAGTGGCGGAGAGT 508
Db 59411 GATCAGCAAGCTGCGCAGCTAGCTGCTTCAAGCACAAGACCGCAAGTGGCGGAGAGT 59352
QY 509 CCTTCAGACGATTTGAAAGCTCTTCATGAGTACGGACCCAGCAGCTTATGTTGCGGT 568
Db 59351 CCTTCAGACGATTTGAAAGCTCTTCATGAGTACGGACCCAGCAGCTTATGTTGCGGT 59292
QY 569 CTATATACCAAGCTTGTGACCTCTCGAGATCCCAAGTATGAGGAGAGCGGC 628
Db 59291 CTATATACCAAGCTTGTGACCTCTCGAGATCCCAAGTATGAGGAGAGCGGC 59232
QY 629 CATCCAAAGCTAGTGAATCTTACAGCATGTAGGGATCGATTGCCCGCAGACCTCG 688
Db 59231 CATCCAAAGCTAGTGAATCTTACAGCATGTAGGGATCGATTGCCCGCAGACCTCG 59172
QY 689 TCGATGACGATGTTCTGCTCGCTGAAATTTGGCTATGTTGAGCAAAAGTTGACCAAGT 748
Db 59171 TCGATGACGATGTTCTGCTCGCTGAAATTTGGCTATGTTGAGCAAAAGTTGACCAAGT 59112
QY 749 CAACAGAGAGGCTACTGCTACCTTCAAGCCCTTAAACAGATTTGCAATGAGTGGG 808
Db 59111 CAACAGAGAGGCTACTGCTACCTTCAAGCCCTTAAACAGATTTGCAATGAGTGGG 59052
QY 809 CTTGACGAGGCGGCAATATTGGGTTGAGGAGCGACCCACAGATGATTAGCGGC 868
Db 59051 CTTGACGAGGCGGCAATATTGGGTTGAGGAGCGACCCACAGATGATTAGCGGC 58992
QY 869 ACTACACTGCGCGCTTTTGTCTATCTGCGGCCCAAAACCCATTTGAAAGATGTACCGG 928
Db 58991 ACTACACTGCGCGCTTTTGTCTATCTGCGGCCCAAAACCCATTTGAAAGATGTACCGG 58932
QY 929 TGATGCGGCGCGGCAATGATCTTTGCAATCTAGCTTTGAGGTGCGCCCAAT 988
Db 58931 TGATGCGGCGCGGCAATGATCTTTGCAATCTAGCTTTGAGGTGCGCCCAAT 58872
QY 989 GAACATCTTCCACGCTAAGAGCATGAGCATCTTACAAGCAAGTACTAGTATCATG 1048
Db 58871 GAACATCTTCCACGCTAAGAGCATGAGCATCTTACAAGCAAGTACTAGTATCATG 58812
QY 1049 TGATTAACACGAGACTGGAGAAAGTGTGATGCTCTCAAGAGATCAGGGCATTTGCT 1108
Db 58811 TGATTAACACGAGACTGGAGAAAGTGTGATGCTCTCAAGAGATCAGGGCATTTGCT 58752
QY 1109 CATCTCGATATCACTACAGCCGAGTTTGTCTGTACAGCTTAAGGAATTTGCTT 1168
Db 58751 CATCTCGATATCACTACAGCCGAGTTTGTCTGTACAGCTTAAGGAATTTGCTT 58652
QY 1169 AAGCTTCTGACATCTTCAAGAGAGAACTACGATCACAGTATCCGGAGCGCTGAT 1228
Db 58651 AAGCTTCTGACATCTTCAAGAGAGAACTACGATCACAGTATCCGGAGCGCTGAT 58612
QY 1229 CACCATTCGCTTACATGTCTAAGACGCTGAGAAATACTAATGCTTCTGTGAGCAT 1288
Db 58611 CACCATTCGCTTACATGTCTAAGACGCTGAGAAATACTAATGCTTCTGTGAGCAT 58572
QY 1289 TTTGGAGACCTGATTAATTTATACAGAACGCGGAAGTCAATTCGCTTCCAC 1348
Db 58571 TTTGGAGACCTGATTAATTTATACAGAACGCGGAAGTCAATTCGCTTCCAC 58512

QY 1349 AATAGCTCTGAATATATCATTAAGTATACATGACACCAAGCTGCTCAGATCTACAC 1408
Db 58511 AATAGCTCTGAATATATCATTAAGTATACATGACACCAAGCTGCTCAGATCTACAC 58452
QY 1409 AGACACTGTGAATCAATCAAAAGTGAAGAGCATTAAGTCCACATGTGTGAGCTGATGAT 1468
Db 58451 AGACACTGTGAATCAATCAAAAGTGAAGAGCATTAAGTCCACATGTGTGAGCTGATGAT 58392
QY 1469 GCTGCTCTTCAAGAGAGTGGCAGACGAAGCGCTTGGAAAAGAAATCCACCTGATTAAGGA 1528
Db 58391 GCTGCTCTTCAAGAGAGTGGCAGACGAAGCGCTTGGAAAAGAAATCCACCTGATTAAGGA 58332
QY 1529 CACCTTAACAAATTCATTTGGCGATGACAGTCCGATGACCGCCATTCCAGATACGC 1588
Db 58331 CACCTTAACAAATTCATTTGGCGATGACAGTCCGATGACCGCCATTCCAGATACGC 58272
QY 1589 CTATTTGGGCTTTCAAGGGGTCACTTTCAGAGCTGGCGGATCAATATATATGAACTATGA 1648
Db 58271 CTATTTGGGCTTTCAAGGGGTCACTTTCAGAGCTGGCGGATCAATATATATGAACTATGA 58212
QY 1649 CATAGCTGCCAGCGCGCATTTAGAAAGGAAAGAGCGCGCGAGAGAGAACTGG 1708
Db 58211 CATAGCTGCCAGCGCGCATTTAGAAAGGAAAGAGCGCGCGAGAGAGAACTGG 58152
QY 1709 TACTGGGACTGGGACTGCACTGAAACGAGACGCACTGTATCCCGCATTTGGCCGAAACAC 1768
Db 58151 TACTGGGACTGGGACTGCACTGAAACGAGACGCACTGTATCCCGCATTTGGCCGAAACAC 58092
QY 1769 TGGAAACCTGCAAAAGCCACAGCTAGATGATGATTCATTTCAAGGGGAGACACTGCGGC 1828
Db 58091 TGGAAACCTGCAAAAGCCACAGCTAGATGATGATTCATTTCAAGGGGAGACACTGCGGC 58032
QY 1829 TGCTCAACGAGCCAAAGTTAGAGCGCAATATATCATATTTCCAGGCAAGAAACCTTT 1888
Db 58031 TGCTCAACGAGCCAAAGTTAGAGCGCAATATATCATATTTCCAGGCAAGAAACCTTT 57972
QY 1889 AGGACCTATATATTCACACCGGATCGATGACAGAGACAGACAGCAGTGTATCATCTGCC 1948
Db 57971 AGGACCTATATATTCACACCGGATCGATGACAGAGACAGACAGCAGTGTATCATCTGCC 57912
QY 1949 CAGACCTGCGCTGAAATTCACATGCGGTGACACAGCTCAACCGCGGATCGGTTAC 2008
Db 57911 CAGACCTGCGCTGAAATTCACATGCGGTGACACAGCTCAACCGCGGATCGGTTAC 57852
QY 2009 ACCACGCGCCCGGAGAGCGGAGAGATCTCACTACAGCCAGATGAGATCCACTC 2068
Db 57851 ACCACGCGCCCGGAGAGCGGAGAGATCTCACTACAGCCAGATGAGATCCACTC 57792
QY 2069 GCCAAGCAAAACTGAGGGATTCAGTACGCTGTGATTTGATATCTATCGTGGGCCAC 2128
Db 57791 GCCAAGCAAAACTGAGGGATTCAGTACGCTGTGATTTGATATCTATCGTGGGCCAC 57732
QY 2129 TGGGCGCATTCACAAAAGGCTCCGGAAATCCCGGAAGCAGCCAGCTCTAGGGAAAC 2188
Db 57731 TGGGCGCATTCACAAAAGGCTCCGGAAATCCCGGAAGCAGCCAGCTCTAGGGAAAC 57672
QY 2189 GAGTCAACTAGATCAGGTGTGCTTGAATGAAACGAGATGTACTTACAGGTGCGGG 2248
Db 57671 GAGTCAACTAGATCAGGTGTGCTTGAATGAAACGAGATGTACTTACAGGTGCGGG 57612
QY 2249 GTCTCGACGTACGCCCGAGAGAAACACCCAGTAAAGCATTCGGCGCGGACGACTGCT 2308
Db 57611 GTCTCGACGTACGCCCGAGAGAAACACCCAGTAAAGCATTCGGCGCGGACGACTGCT 57552
QY 2309 GCGCGCAATCCCGTGAAGAGCAATACATTTAGCGGTGCGAGATGAGACAGCAGACTA 2368
Db 57551 GCGCGCAATCCCGTGAAGAGCAATACATTTAGCGGTGCGAGATGAGACAGCAGACTA 57492
QY 2369 TGTTCGCGGAGCTACATGCGCAGCGCGGAAATCGGATGGGTTAGAAAGCTCATGGAAGC 2428
Db 57491 TGTTCGCGGAGCTACATGCGCAGCGCGGAAATCGGATGGGTTAGAAAGCTCATGGAAGC 57432
QY 2429 TGATGAGTGTATGATCATGACCTCCAGGCGCAAGTTCTGTGTTCAGAAACGATCTTTGA 2488

Db	57431	TCATGAGTCTGATGACATTCGACTCCGAGCCAGTTCTGTGTGTCAGAACGATCTTTTGA	57372
OY	2489	CTCCAGCTACATAGAGGTAAATAATCGAACTACTCACTTAGCGGAGCCAAACCCGCTT	2548
Db	57371	CTCCAGCTACATAGAGGTAAATAATCGAACTACTCACTTAGCGGAGCCAAACCCGCTT	57312
OY	2549	GGACTGAGAGACCGGAGGGGGCCCATTTGACGACATCGAGACGATTAATTCAGTTCTGGGC	2608
Db	57311	GGACTGAGAGACCGGAGGGGGCCCATTTGACGACATCGAGACGATTAATTCAGTTCTGGGC	57252
OY	2609	ATCGAGCATTTGGTCTGAAAAGGAAGATGGCTGATCGCTCCACACGTAATCGGCCGA	2668
Db	57251	ATCGAGCATTTGGTCTGAAAAGGAAGATGGCTGATCGCTCCACACGTAATCGGCCGA	57192
OY	2669	TGGAAAGAGACTCACTCAGCAGCACTTAAATGCGTTCTGACATGTTCCGCAAGATGTT	2728
Db	57191	TGGAAAGAGACTCACTCAGCAGCACTTAAATGCGTTCTGACATGTTCCGCAAGATGTT	57132
OY	2729	CATGATATCACACACCAAGGTGTAATCTACTGTTTTGGACACGGTCACCGAATGATTTCT	2788
Db	57131	CATGATATCACACACCAAGGTGTAATCTACTGTTTTGGACACGGTCACCGAATGATTTCT	57072
OY	2789	GGTTCATGGGAAGAAACTTACAGAAATGGTGCTTCAATCTGTTGACGGGGTGTTCAA	2848
Db	57071	GGTTCATGGGAAGAAACTTACAGAAATGGTGCTTCAATCTGTTGACGGGGTGTTCAA	57015
OY	2849	CAAAATGGGACCCGATCTCAATTCGATGACAGACAGATTTGAAAGCGCTACAGGT	2908
Db	57014	CAAACTGGGACCCGATCTCAATTCGATGACAGACAGATTTGAAAGCGCTACAGGT	56955
OY	2909	GGTCCACGAATATTTCCCAAGAGTTGACGCTTAAAGACTAATTAAGATCATATCGGA	2968
Db	56954	GGTCCACGAATATTTCCCAAGAGTTGACGCTTAAAGACTAATTAAGATCATATCGGA	56895
OY	2969	TTCTACTCAAAAGCCCACTACCAAGACGGCATATGCGGCTTCTTAAGGATCT	3028
Db	56894	TTCTACTCAAAAGCCCACTACCAAGACGGCATATGCGGCTTCTTAAGGATCT	56835
OY	3029	GGCCAAATACATCTGTAAAGACGCACTTTCCAGGACACAGAGCCAGGCTCGAGACG	3088
Db	56834	GGCCAAATACATCTGTAAAGACGCACTTTCCAGGACACAGAGCCAGGCTCGAGACG	56775
OY	3089	GACGGTCTTAAAGCTGGCCAGCTGGCGGCGGATCAGAAGTCGATGAGACTGCGCTCCA	3148
Db	56774	GACGGTCTTAAAGCTGGCCAGCTGGCGGCGGATCAGAAGTCGATGAGACTGCGCTCCA	56715
OY	3149	GGCCAGGAGCTGCTTAGTGCCCTGTATTAACCTGTAATACCCCGCAATGACCTTTTAAT	3208
Db	56714	GGCCAGGAGCTGCTTAGTGCCCTGTATTAACCTGTAATACCCCGCAATGACCTTTTAAT	56655
OY	3209	GGCCGAGCTGGCCAAAGATATATGAGACTGTGCGCCGATTCGTCGATTCGACATGAG	3268
Db	56654	GGCCGAGCTGGCCAAAGATATATGAGACTGTGCGCCGATTCGTCGATTCGACATGAG	56595
OY	3269	GGCGCAAAAGCTGCAATTCGAGTGGCAATTCGACTAGTACTCTCCATTGAGCAG	3328
Db	56594	GGCGCAAAAGCTGCAATTCGAGTGGCAATTCGACTAGTACTCTCCATTGAGCAG	56535
OY	3329	TAGCACTCCCAAGCTTTTGCAAGTCCCTGTGGGTCCATTTGCTTGCTTCAGAGCCA	3388
Db	56534	TAGCACTCCCAAGCTTTTGCAAGTCCCTGTGGGTCCATTTGCTTGCTTCAGAGCCA	56475
OY	3389	GCACCAACCAATTAGATCAGCTCTACTACGACGCTCCCGGAGAGCTTCCGGGAGCA	3448
Db	56474	GCACCAACCAATTAGATCAGCTCTACTACGACGCTCCCGGAGAGCTTCCGGGAGCA	56415
OY	3449	GGAGCTGCTCTTTCTCTCGAGCTGACATTCAGACACAACTTCAGAAAGAGCTGAGAGA	3508
Db	56414	GGAGCTGCTCTTTCTCTCGAGCTGACATTCAGACACAACTTCAGAAAGAGCTGAGAGA	56355
OY	3509	GATCCGGCACTGCTTGGCGGTGAGTACAGACGGCGCTGGCGCCCAATGGCTTCAATGG	3568

Db	56354	GATTCGGCACTGCTTGGCGGTAGTACAGACGGCGCTGGCGCCAAATGGCTTCAATGG	56295
Qy	3569	ACACTTGAGTATCAAGATCAGGGCCAAACGAAATTCGTGTGCATCCCTGTCCAACTC	3628
Db	56294	ACACTTCAGTATCAAGATCAGGGCCAAACGAAATTCGTGTGCATCCCTGTCCAACTC	56239
Qy	3629	CAAGCGCAATGCTGGCCAAACCTACCCGTAATAATACCTGATCAGAACAAATGAG	3688
Db	56234	CAAGACGCAATGCTGGCCAAACCTACCCGTAATAATACCTGATCAGAACAAATGAG	56175
Qy	3689	GCTGATTAATCTGSAACGGGAAAGACAACCTCAAGACGCCAAGTCAACCACTGACATG	3748
Db	56174	GCTGATTAATCTGSAACGGGAAAGACAACCTCAAGACGCCAAGTCAACCACTGACATG	56115
Qy	3749	CAGGTGATCAGCGTCTCGATAAATATGCTGAATAATGAGAGCTGATCTGCGACGA	3808
Db	56114	CAGGTGATCAGCGTCTCGATAAATATGCTGAATAATGAGAGCTGATCTGCGACGA	56055
Qy	3809	CCTGATGAGACGGAAGTGTGCTGTGGCTTGACGCTTAACAAAGATCAGCCCGTGA	3868
Db	56054	CCTGATGAGACGGAAGTGTGCTGTGGCTTGACGCTTAACAAAGATCAGCCCGTGA	55995
Qy	3869	GTTGCTTCAGACGTCACTACTAATCCTGGGGAATTTGATCAAGGGCGGAAACTGTGACT	3928
Db	55994	GTTGCTTCAGACGTCACTACTAATCCTGGGGAATTTGATCAAGGGCGGAAACTGTGACT	55935
Qy	3929	GCCCAATAGCACTTTAGATCATCATGCAATCTGCTTAACTTGTGAGCGGAGCA	3988
Db	55934	GCCCAATAGCACTTTAGATCATCATGCAATCTGCTTAACTTGTGAGCGGAGCA	55875
Qy	3989	TACGACGTGTTCATCTGTGGCTGCACTGTCTCAATTAATTTAGAGCAACAAAT	4048
Db	55874	TACGACGTGTTCATCTGTGGCTGCACTGTCTCAATTAATTTAGAGCAACAAAT	55815
Qy	4049	GGGTCAACACTGAGTGCATTTCTAGAGGATTTTGTCTGAAGTCACTCCAGTCTATCA	4108
Db	55814	GGGTCAACACTGAGTGCATTTCTAGAGGATTTTGTCTGAAGTCACTCCAGTCTATCA	55755
Qy	4109	ACACAGCAAGAGGCTTTGCGGATATGCACTGATGATACCAAGATAGCACCATCTT	4168
Db	55754	ACACAGCAAGAGGCTTTGCGGATATGCACTGATGATACCAAGATAGCACCATCTT	55695
Qy	4169	GCCTCTGATCTGTTCATCAACTGTCAATCCGTTAATTGCGACGGGTAGTTCCAC	4228
Db	55694	GCCTCTGATCTGTTCATCAACTGTCAATCCGTTAATTGCGACGGGTAGTTCCAC	55635
Qy	4229	AAATTTGGCGCATCAAGATCCTGTGGAAGGACGGAACACATGTTGAGATAC	4288
Db	55634	AAATTTGGCGCATCAAGATCCTGTGGAAGGACGGAACACATGTTGAGATAC	55575
Qy	4289	AGACGCCCACTTGACATTTGTTCCTCCCAATCTGGCGCATCTGGCGACGACACGCAATC	4348
Db	55574	AGACGCCCACTTGACATTTGTTCCTCCCAATCTGGCGCATCTGGCGACGACACGCAATC	55515
Qy	4349	GATGTTGCAAAAGCTGGGCTTCTGCATCTCAAGCTGACTTTGTTCTGGCGAAGA	4408
Db	55514	GATGTTGCAAAAGCTGGGCTTCTGCATCTCAAGCTGACTTTGTTCTGGCGAAGA	55455
Qy	4409	GAAAGTCAAGCCGGAAGCTGTCACTGTCTAATCCACCAAGTTAAGGCTCTCAACGTTGA	4468
Db	55454	GAAAGTCAAGCCGGAAGCTGTCACTGTCTAATCCACCAAGTTAAGGCTCTCAACGTTGA	55395
Qy	4469	CATCAGAGAAGCAGCGGAACCTGATCAGTGGGGAGGAAGCTCTCAAGAAGAACTCTCGC	4528
Db	55394	CATCAGAGAAGCAGCGGAACCTGATCAGTGGGGAGGAAGCTCTCAAGAAGAACTCTCGC	55335
Qy	4529	GGCATCTGTCTATGATTTGCGGAGACCTTTAATAGAAATTCCTGC - TCTGTCACCAAA	4586
Db	55334	GGCATCTGTCTATGATTTGCGGAGACCTTTAATAGAAATTCCTGCATCAGACCAACCAAA	55275
Qy	4587	ACAAGACACAGACGGGTGCTTCCCTCGGCTGAGAAGAAATGAGAGACGGCGGACAT	4646
Db	55274	ACAAGACACAGACGGGTGCTTCCCTCGGCTGAGAAGAAATGAGAGACGGCGGACAT	55215

OY	4647	AAATTAATATATTTTACATTAACATAATTATTTAATCAATTAATTAACGATCATCCGT	4706
Dd	55214	TAAATTAATATTTATTTACATTAACATAATTATTTAATCAATTAATTAACGATCATCCGT	55155S
OY	4707	ACTTGCTGAATCGCTGCATGGATATATTTCAGCAGATGCCGCCGAAAAGAAAGTC	4766
Dd	55154	ACTTGCTGAATCGCTGCATGGATATATTTCAGCAGATGCCAATGCCGCCGAAAAGAAAGTC	55095S
OY	4767	AAGGGCCATGCCCATCTAATGTGAGCAGAAAACTATTATACATAAACGGAACTAAA	4826
Dd	55094	AAGGGCCATGCCCATCTAATGTGAGCAGAAAACTATTATACATAAACGGAACTAAA	55035S
OY	4827	GCGAGTAATTCGCAAAAGTGACTAAATTTGTAGCCAATCCGCAATCCTCTTCAC	4886
Dd	55034	GCGAGTAATTCGCAAAAGTGACTAAATTTGTAGCCAATCCGCAATCCTCTTCAC	54975S
OY	4887	ATCAGTGTACGCTTTTTTACACAGCTCPAATTAAGAAGCTATTCTGTTTAATTGTA	4946
Dd	54974	ATCAGTGTACGCTTTTTTACACAGCTCPAATTAAGAAGCTATTCTGTTTAATTGTA	54915S
OY	4947	AGCCTATATACCGCTCTTATNGAAACTACAGCGTGTGCTTTCTACCTTGTTCAGGC	5006
Dd	54914	AGCCTATATACCGCTCTTATNGAAACTACAGCGTGTGCTTTCTACCTTGTTCAGGC	54855S
OY	5007	GACTTTGTATCCTATTATTCGTTTCGCTCAACTAAATTCCTAATTACTATTATTAATACA	5066
Dd	54854	GACTTTGTATCCTATTATTCGTTTCGCTCAACTAAATTCCTAATTACTATTATTAATACA	54795S
OY	5067	TTCTGCTCTTACAACTGAATCTATTTTGTATTAATTAATTAATACAGAGCCACAAAG	5126
Dd	54794	TTCTGCTCTTACAACTGAATCTATTTTGTATTAATTAATTAATTAATACAGAGCCACAAAG	54735S
OY	5127	GATTAATAATCAATAAAAAA	5145
Dd	54734	GATTAATAATCAATAAAAAA	54716

[illegible]

REFERENCE AUTHORS TITLE JOURNAL	Direct Submission Unpublished 3 (bases 1 to 182960) Morley,K.C. Direct Submission Submitted (11-SEP-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 4 (bases 1 to 182960) BCM-HGSC.
REFERENCE AUTHORS TITLE JOURNAL	Direct Submission Submitted (11-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 5 (bases 1 to 182960) BCM-HGSC.
REFERENCE AUTHORS TITLE JOURNAL	Direct Submission Submitted (12-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 6 (bases 1 to 182960) BCM-HGSC.
REFERENCE AUTHORS TITLE JOURNAL	Direct Submission Submitted (18-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA On Jun 18, 2002 this sequence version replaced gi:18129369.
COMMENT	

Baruch, P., Burkett, C., Butrell, K.L., Byrd, N.C., Cartron, T.F.,
Carrion, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
Davy, Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
Dem, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,
Dugan-Rocha, S., Durbin, K.J., Eamhart, C., Edgar, D., Edwards, C.C.,
Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,
Foster, P., Frantz, P., Gablis, A., Gao, J., Garcia, A., Garner, T.,
Garra, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,
Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A.,
Hernandez, J., Hernandez, O., Hodgson, A., Hoques, N., Holloway, C.,
Hollins, B., Homsi, F., Howard, S., Huber, J., Hultzy, S., Hume, J.,
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
Kovar, C., Kratovic, J., Kurehli, A., Landry, N., Leal, B., Lewis, L.C.,
Lewis, L., Li, U., Li, Z., Lichtharge, O., Lieu, C., Liu, J., Liu, W.,
Loubeged, H., Lozada, R.M., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Ma, J., Maheshwari, M., Mapua, P., Martin, R., Matlindale, A.,
Martinez, E., Massey, E., Mawlinery, E., Meleod, M.P., Meador, K.,
Mei, G., Mezker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
Morgan, M., Morris, S., Moser, N., Neal, J., Newton, J., Newton, N.,
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S.,
Ogun, M., Okunodu, G., Oragunye, N., Ortedo, R., Pace, A., Rayton, B.,
Ogutu, J., Perez, L., Peters, L., Pickens, R., Primus, B., Pu, L.L.,
Peetley, J., Ren, Y., Rivers, M., Rojas, A., Rojudoakan, I., Rolle, M.,
Quiles, M., Saevry, G., Scherer, S., Scott, G., Shen, H., Shoostart, N.,
Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H.,
Stone, H., Sutton, A., Svatek, A., Tabori, P., Tameris, A., Tamerisa, K.,
Thomas, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
Thomas, S., Umanli, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,
Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
Watlington, S., Williams, G., Williamson, A., Wiczynski, R., Wooden, S.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.

Direct Submission

Unpublished

3 (bases 1 to 182960)

Worley, K.C.

Direct Submission

Submitted (11-SEP-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 182960)

BCM-HGSC

Direct Submission

Submitted (11-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

5 (bases 1 to 182960)

BCM-HGSC

Direct Submission

Submitted (12-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

6 (bases 1 to 182960)

BCM-HGSC

Direct Submission

Submitted (18-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Jun 18 2002 this sequence version replaced gi:18129369.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.

Overlapping clones are noted at the beginning and end of the
features listing.

ANNOTATION OF FEATURES:

gc-helpp@bcm.tmc.edu

STS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.bgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES
source
1. 182960
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3L"
/clone="RP98-2C22"
39171..48280
misc_feature
/note="polymorphic rpo element which does not exist in celera data and in DMR"
/function="polymorphic site"

ORIGIN

Query Match 98.2% Score 5051.8; DB 2; Length 182960;

Best Local Similarity 99.6%; Pred. No. 0; Mismatches 17; Indels 5; Gaps 3;

Matches 5097; Conservative 0; Mismatches 17; Indels 5; Gaps 3;

29 GTTCCGCGAGACGAGGACGAGATGACATTCGACATGCGCTATCCGAGCCAGCGA 88
176386 GTTCCGCGAGACGAGGACGAGATGACATTCGACATGCGCTATCCGAGCCAGCGA 176327
89 CTTGATGAGCTTTATCCAGCAATATGCGGCGGACATGCGTGTGAAGTACAGCTGC 148
176326 CTTGATGAGCTTTATCCAGCAATATGCGGCGGACATGCGTGTGAAGTACAGCTGC 176267
149 CGAGATCTGCTGATCTCTTACGAGGACGAGCAATGCTGCTGCGAGGACATGGG 208
176266 CGAGATCTGCTGATCTCTTACGAGGACGAGCAATGCTGCTGCGAGGACATGGG 176207
209 ATTCTTATTCGAGCTTGTGATGCGATGCGGCGGACGACCTTTAAGTTGACAAA 268
176206 ATTCTTATTCGAGCTTGTGATGCGGCGGACGACCTTTAAGTTGACAAA 176147
269 GTTCCGCGAGGCTTCTGAGGCTATTAAGCGATGGGCGGACGATTTAATGCATAC 328
176146 GTTCCGCGAGGCTTCTGAGGCTATTAAGCGATGGGCGGACGATTTAATGCATAC 176087
329 GGCTACGCTTTCGACATGATGATCGGCTGGGAGAGGAGGAGGACAGTCCGCGA 388
176086 GGCTACGCTTTCGACATGATGATCGGCTGGGAGAGGAGGAGGACAGTCCGCGA 176027
389 GAAGCGCACTTCTGCTGCGGACCTTCATGAGGACAGATGCTTCCGCCAGCGCT 448
176026 GAAGCGCACTTCTGCTGCGGACCTTCATGAGGACAGATGCTTCCGCCAGCGCT 175967
449 GATGACAGAGTGGCGCACTAGCTGTTACGACAGAACCCAGGTCGCGAGAGTT 508
175967 GATGACAGAGTGGCGCACTAGCTGTTACGACAGAACCCAGGTCGCGAGAGTT 174827

175966 GATGACAGAGTGGCGCACTAGCTGTTACGACAGAACCCAGGTCGCGAGAGTT 175907
509 CTTTCAGACGATTTGAGACGCTTTTCATGATGAGGACGAGGACGCTTATGTTCCGCT 568
175906 CTTTCAGACGATTTGAGACGCTTTTCATGATGAGGACGAGGACGCTTATGTTCCGCT 175847
569 CTATATACACAGTTTGTGACCTTTCGAGGATCCCACTATTAATGAGGAGGCGGC 628
175846 CTATATACACAGTTTGTGACCTTTCGAGGATCCCACTATTAATGAGGAGGCGGC 175787
629 CATCAACGCTATGGAATCTCAAGCATGTGAGGAGATGATGGCCGACGCTCCG 688
175786 CATCAACGCTATGGAATCTCAAGCATGTGAGGAGATGATGGCCGACGCTCCG 175727
689 TCGCATGAGACGATTTCTGCTCGGATGGAATTTGAGGACGAGGATTTGAGGACG 748
175726 TCGCATGAGACGATTTCTGCTCGGATGGAATTTGAGGACGAGGATTTGAGGACG 175667
749 CAACAGAGGCTTACTGCTACTCTTCAAGCCCTTTAAACAGAAATGAGATGAGTGG 808
175666 CAACAGAGGCTTACTGCTACTCTTCAAGCCCTTTAAACAGAAATGAGATGAGTGG 175607
809 CTTGAGAGGCGGACCAATTTGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 868
175606 CTTGAGAGGCGGACCAATTTGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 175547
869 ACTACATCGGCGCTTTCATGATGATGCGGCGGACCAAGCCCAATGAGATGAGTGG 928
175546 ACTACATCGGCGCTTTCATGATGATGCGGCGGACCAAGCCCAATGAGATGAGTGG 175487
929 TGATGCGGCGGCGGTAACATGAGATTTTGAATCTTGAAGTGTGCTCCGCAAT 988
175486 TGATGCGGCGGCGGTAACATGAGATTTTGAATCTTGAAGTGTGCTCCGCAAT 175427
989 GAACATTTCCAGCTTACAGGACATGAGGATGATGATGAGGATGATGATGATGATG 1048
175426 GAACATTTCCAGCTTACAGGACATGAGGATGATGATGAGGATGATGATGATGATG 175367
1049 TGATTAACAGCAGATGAGGAGAAAGTGTGATGCTCTCAAGAAATCAGGAGCTTCT 1108
175366 TGATTAACAGCAGATGAGGAGAAAGTGTGATGCTCTCAAGAAATCAGGAGCTTCT 175307
1109 CATTCAGCTATCACTCAGCGCGAGTTTGTGCTGATGATGATGATGATGATGATG 1168
175306 CATTCAGCTATCACTCAGCGCGAGTTTGTGCTGATGATGATGATGATGATGATG 175247
1169 AAGCTTGTGATGATCTCTCAAGGAGAACTACATGATGATGATGATGATGATGATG 1228
175246 AAGCTTGTGATGATCTCTCAAGGAGAACTACATGATGATGATGATGATGATGATG 175187
1229 CACCATGCTCACTATGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1288
175186 CACCATGCTCACTATGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 175127
1289 TTGAGACACCTGATTAATTAATTAACAGAACGCGGAGGATGATGATGATGATGAT 1348
175126 TTGAGACACCTGATTAATTAATTAATTAACAGAACGCGGAGGATGATGATGATGAT 175067
1349 AATAGCTCTGAAGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1408
175066 AATAGCTCTGAAGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 175007
1409 AGACACTGATATCAATCAAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1468
175006 AGACACTGATATCAATCAAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 174947
1469 GCTGCTCTTGAAGAGTGCAGACGAAAGGCTTGAAGGAGGAGGAGGAGGAGGAGG 1528
174946 GCTGCTCTTGAAGAGTGCAGACGAAAGGCTTGAAGGAGGAGGAGGAGGAGGAGG 174887
1529 CACCTTAACAAATTCATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1588
174886 CACCTTAACAAATTCATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 174827

Db	172669	CAGGAGTACGCGTCTCGATTAAATATAGCTGAAAAAGAGACTGATACCTGGCGACGAA	172610
QY	3809	CCTGATGAGAGCGAAAGTGGTGGCTGTGGCTTTGACGCTAACAAAGGATCAGCCCTGCGA	3868
Db	172609	CCGTATGAGAGCGGAAGTGGTGGCTGTGGCTTTGACGCTAACAAAGGATCAGCCCTGCGA	172550
QY	3869	GTTCGTTCCAGCGTCACTTAACCTGGGGGATTTTGATCAAGGGCGGAAACTGTAGCT	3928
Db	172549	GTTCGTTCCAGCGTCACTTAACCTGGGGGATTTTGATCAAGGGCGGAAACTGTAGCT	172490
QY	3929	GGCCATAAGACACTTATAGATCGATCATGTGCGAAATGCTGCTTAACATTTCTGGAGCGGAGCA	3988
Db	172489	GGCCATAAGACACTTATAGATCGATCATGTGCGAAATGCTGCTTAACATTTCTGGAGCGGAGCA	172430
QY	3989	TACGACGTGGTCATGCTGGCGCTGACGCTCAGTAAGATTATGAGAGCAACAAAT	4048
Db	172429	TACGACGTGGTCATGCTGGCGCTGACGCTCAGTAAGATTATGAGAGCAACAAAT	172370
QY	4049	GGGTCAACAACGTGATGCATTTTAAAGCTGATTTTGTCTGAAATCATCCAGTGTATCA	4108
Db	172369	GGGTCAACAACGTGATGCATTTTAAAGCTGATTTTGTCTGAAATCATCCAGTGTATCA	172310
QY	4109	ACACAGCAAGAGGCGCTTGGGGGATATPCAGCTGATGATPCCAAGATAGCACATCTT	4168
Db	172309	ACACAGCAAGAGGCGCTTGGGGGATATPCAGCTGATGATPCCAAGATAGCACATCTT	172250
QY	4169	GGCTCTGATCTGTCCATCAACATTTGTCATCCGGTAATTTGCGACGGGTGATTTCCAC	4228
Db	172249	GGCTCTGATCTGTCCATCAACATTTGTCATCCGGTAATTTGCGACGGGTGATTTCCAC	172190
QY	4229	AAATTTGTGCGCCATCAAGATCTCTGGAAGTGACCGAACACCATGTGTTCCGAGATCAC	4288
Db	172189	AAATTTGTGCGCCATCAAGATCTCTGGAAGTGACCGAACACCATGTGTTCCGAGATCAC	172130
QY	4289	AGACCCCACTTGGACATTTGTTCCTCCCAATCTGGGGCGAATCGGCGGACGACCGCAATC	4348
Db	172129	AGACCCCACTTGGACATTTGTTCCTCCCAATCTGGGGCGAATCGGCGGACGACCGCAATC	172070
QY	4349	GATGTTCCGCAAGCTGCGGCTCTTCTGCAATCGTCAAGCTGTATTTGTTCTGGGCGAGGA	4408
Db	172069	GATGTTCCGCAAGCTGCGGCTCTTCTGCAATCGTCAAGCTGTATTTGTTCTGGGCGAGGA	172010
QY	4409	GAAAGTCAAGCCGAAGCTGTCAAGTGTAAATCCACGAAGTTAGGCTCTCAACGCTGA	4468
Db	172009	GAAAGTCAAGCCGAAGCTGTCAAGTGTAAATCCACGAAGTTAGGCTCTCAACGCTGA	171950
QY	4469	CATCGAAGACGCGGAACTGCATCAGTGGGGGAGGAAGCTCTACAAAGAACCTCTCCGC	4528
Db	171949	CATCGAAGACGCGGAACTGCATCAGTGGGGGAGGAAGCTCTACAAAGAACCTCTCCGC	171890
QY	4529	GGCATCGTGTATGATTTGCGGGAGGCCCTTAATAGGATTCCTGCG--TCGGCCACACAA	4586
Db	171889	GGCATCGTGTATGATTTGCGGGAGGCCCTTAATAGGATTCCTGCGATCAGACACACAA	171830
QY	4589	ACAAAGACAGACGCGGTTCTTCCCTCGGCTGAGAAAGAAAGTGAAGACGCGCGACAT	4646
Db	171829	ACAAAGACAGACGCGGTTCTTCCCTCGGCTGAGAAAGAAAGTGAAGACGCGCGACAT	171770
QY	4647	TAAATTAATATTAATTAACTTAATCTTAATTAATTAATTAATTAATTAATTAATTAATTA	4706
Db	171769	TAAATTAATATTAATTAACTTAATCTTAATTAATTAATTAATTAATTAATTAATTAATTA	171710
QY	4707	ACTTGTGTAACGTCTGATGCATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	4766
Db	171709	ACTTGTGTAACGTCTGATGCATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	171650
QY	4767	AAGGCGCATCGCCCATCTATATGTAGACAGAAAAAATATTATATACATAAACGGGAACTAA	4826
Db	171649	AAGGCGCATCGCCCATCTATATGTAGACAGAAAAAATATTATTAATTAATTAATTAATTA	171590
QY	4827	GGAGTAAATCCGCAAGTGTACTAAATTTGATGCGCACTCCGCAATCTCACTTCTCAC	4886

D6	171869	CCGAGTAATCCCGGAAGTGAACATAAATTGAGCAACTCCGCAATCCTCACTTCAC	171551
OY	4887	ATCATGTGTACGCTCTTTTACCAGCTCTTAACATATTAACGGTGATTCGTTTAATTGTA	4946
D6	171529	ATCATGTGTACGCTCTTTTACCAGCTCTTAACATATTAACGGTGATTCGTTTAATTGTA	171470
OY	4947	AGCCTATATACCGGCTCTTATATGAAACGCACAACGGTGCTTTTACCTTTTTCGAGGC	5006
D6	171469	AGCCTATATACCGGCTCTTATATGAAACGCACAACGGTGCTTTTTCACCTTTTTCGAGGC	171410
OY	5007	GACCTTTGATCCTTATTTTCGTTTCGCGTCACACTAAATTCATATTCATTTATTAATACGA	5066
D6	171409	GACCTTTGATCCTTATTTTCGTTTCGCGTCACACTAAATTCATATTCATTTATTAATACGA	171350
OY	5067	TTCGCTCTTACAACTGAACTATTTTGTATTAATTAATTAATACCAAGCCACACAAG	5126
D6	171349	TTCGCTCTTACAACTGAACTATTTTGTATTAATTAATTAATTAATACCAAGCCACACAAG	171290
OY	5127	GATTAATAATCAATATAAAA	5145
D6	171289	GATTAATAATCAATATAAAA	171271
RESULT 9			
AEO05593		313157 bp DNA linear INV 11-AUG-2005	
LOCUS		Drosophila melanogaster chromosome 3L, section 74 of 83 of the complete sequence.	
DEFINITION		AEO03593 AEO02647 AE014296	
ACCESSION		AEO03593.2 GI:23094195	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM		Drosophila melanogaster (fruit fly)	
REFERENCE		Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 313157) Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,Y.H., Blazer,R.G., Champe,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabot,D., Anil,I.J.F., Agdayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Benos,P.V., Bernier,B.P., Bhandari,D., Bolshakov,S., Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brotlier,P., Burks,K.C., Buam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,T., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de Pablos,B., Delcher,A., Deng,Z., May,A.D., Dew,I., Dietz,S.M., Dodson,K., Doug,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.A., Ferraz,C., Periera,S., Fleischmann,W., Fosler,C., Gabrielian,A.E., Gang,N.S., Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Helman,T.J., Hernandez,J.R., Houck,J., Hoctin,D., Houston,K.A., Howland,T.J., Wei,M.H., Iqbalwala,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z., Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kul,D., Lai,Z., LaRoche,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Matcel,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Morris,J.T., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nuskey,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S., Pollard,J., Puti,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Stamps,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stepleton,M., Strong,R., Sun,E., Svrtenski,R., Tecor,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassenaar,D.A., Weinstock,G.M., Weissbach,J., Williams,S.M., Woodgett, Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F., Zaveri,U.S., Zhan,M., Zhang,Q., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,	


```

Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.
The genome sequence of Drosophila melanogaster
JOURNAL
PUBMED
10731132
2 (bases 1 to 313157)
Celniker, S.E., Wheeler, D.A., Krommiller, B., Carlson, J.W.,
Halpern, A., Patel, S., Adams, M., Champe, M., Dugan, S.P., Frise, E.,
Hodgson, A., George, R.A., Hoskins, R.A., Laverly, T., Muzny, D.M.,
Neelson, C.R., Pacle, J.M., Park, S., Pfeiffer, B.D., Richards, S.,
Sodergren, E.J., Svirskae, R., Tabor, P.E., Wan, K., Stapleton, M.,
Sutton, G.G., Venter, C., Weinstock, G., Scherer, S.E., Myers, E.W.,
Gibbs, R.A. and Rubin, G.M.
Finishing a whole-genome shotgun: release 3 of the Drosophila
melanogaster euchromatic genome sequence
JOURNAL
PUBMED
12537568
3 (bases 1 to 313157)
Mungall, C.J., Matthews, B.B., Campbell, K.S.,
Miera, S., Crosby, M.A., Kaminker, J.S., Milburn, G.H., Prochuk, S.E.,
Hirdecky, P., Huang, Y., Kaminker, J.S., Milburn, G.H., Prochuk, S.E.,
Smith, C.D., Tupy, J.L., Whitfield, E.J., Bayraktaroglu, L.,
Berman, B.P., Bertencourt, B.R., Celniker, S.E., de Grey, A.D.,
Drysdale, R.A., Harris, N.L., Richter, J., Russo, S., Schroeder, A.J.,
Shu, S.Q., Stapleton, M., Yamada, C., Ashburner, M., Gelbart, W.M.,
Rubin, G.M. and Lewis, S.E.
Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review
JOURNAL
PUBMED
12537572
4 (bases 1 to 313157)
Kaminker, J.S., Bergman, C.M., Krommiller, B., Carlson, J.,
Svirskae, R., Patel, S., Frise, E., Wheeler, D.A., Lewis, S.E.,
Rubin, G.M., Ashburner, M. and Celniker, S.E.
The transposable elements of the Drosophila melanogaster
euchromatin: a genomic perspective
JOURNAL
PUBMED
12537573
5 (bases 1 to 313157)
Celniker, S., Carlson, J., Wan, K., Pfeiffer, B., Frise, E., George, R.,
Hoskins, R., Stapleton, M., Pacle, J., Park, S., Svirskae, R.,
Smith, E., Yu, C. and Rubin, G.
Berkeley Drosophila Genome Project
Unpublished
6 (bases 1 to 313157)
Adams, M.D., Celniker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
7 (bases 1 to 313157)
FlyBase
Direct Submission
Submitted (22-JUL-2005) FlyBase, Harvard University, Biological
Laboratories, 16 Divinity Ave, Cambridge, MA 02138, USA
On Sep 17, 2002 this sequence version replaced g1:7296371.
location/Qualifiers
1. 313157
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/chromosome="3L"
/complement(<3239..>4179)
/gene="CG10586"
/locus_tag="Dmel CG10586"
/old_locus_tag="CG10586"
/map="78A1-78A1"
/db_xref="FLYBASE:FBgn0037036"
/complement(join(<3239..3462,3576..>4179))
/gene="CG10586"
/locus_tag="Dmel CG10586"
/old_locus_tag="CG10586"
/product="CG10586-RA"
/db_xref="FLYBASE:FBgn0037036"

CDS
complement(join(3239..3462,3576..4179))
/gene="CG10586"
/locus_tag="Dmel CG10586"
/old_locus_tag="CG10586"
/note="CG10586 gene product from transcript CG10586-RA"
/codon_start=1
/product="CG10586-PA"
/protein_id="AA051660.1"
/db_xref="FLYBASE:FBgn0037036"
/translacion="MKRLFLFLAGLILNNHALQHNETIDLKAKIVLPAPAYQTV
IGRVTNNATLGGVILVAMRYENNFCGGTILHEILVITAAHCEDEARAKAWSDGCI
SRSEKIGRQKRVGRFTKSAQFKNVTMMDAVILNRPNGKINGTSLSTLTPCO
TMVSGKNTNPDDEGKMLRTVSPVIEKRICREKRSYSISDSMTFCASVIGKMD
ACTYDSGGLPVYKQVGVFSIGCASRRYPGVYDVHVKPTVAGTALISRRR"
<4431..>7607
/gene="CG10588"
/locus_tag="Dmel CG10588"
/old_locus_tag="CG10588"
/map="78A1-78A1"
/db_xref="FLYBASE:FBgn0037037"
<4431..>7607
/gene="CG10588"
/locus_tag="Dmel CG10588"
/old_locus_tag="CG10588"
/product="CG10588-RA"
/db_xref="FLYBASE:FBgn0037037"
4431..7607
/gene="CG10588"
/locus_tag="Dmel CG10588"
/old_locus_tag="CG10588"
/note="CG10588 gene product from transcript CG10588-RA"
/codon_start=1
/product="CG10588-PA"
/protein_id="AA051661.2"
/db_xref="FLYBASE:FBgn0037037"
/translacion="MCFRPLGRRITRKQVQKALNROQYPMSTASVYVVOYHEPDKSD
GDSKLYALTLNSGLRAMLISDSYIDPSSIHRSRSLNSTENSTENHATLACAYLVG
GSFSEPOOYGLAHFVEHMIFMGSEKPEPVNEPDSFPTKSGGPNHATNEETCFYE
LDOTHLDRGMDLPMNLKAPLMLPDAMSRRSAVQSFQDTHMDEVARQDILSLAS
EGYPHGFSNGKNTKTLKEGVDSLSLHEKELFYDHRGSRMVAALQAOISLDEEL
LVHRCADIPISQONSIDVSOLOKARPOFVDVPLVQVDEDSCKELTMVLPMPKN
FYRSKPMFTISQILIGVGSCLAYLRHHMCISVAVGVAESSPDSNISYLENICV
LSDDGFPHIDEVLEATFAWKLINSQLODYSQIESQIENNNRFOIIPPIDHVO
IVESFNILPSKDYLTGPOLFOYVESAIELRQHNKFNFNIMISSYIPEKNYDOR
EPMNGQFKTISMPKQTMWEOGATIKELAHYPOPNPFTTDFKIHVESGKPHVSR
PKALIRNDICEIMERODNIFKLPDGINLFIPLVSVQYWLGVLFYLVFRRMA
EQLYPALEAGLTGYIGDKGLVNRVSGVNEKPLVAYEILNNMOTIELDGOVNAFK
DLKRQITVNALINGSLNDLRSLILENKRFSMTSKYESVDDITMDIKSKFENPHK
MYVKGILQGNFTEDQATDLMQKVLDTYKSEKLDLSALDNLQIPLGSIYLRKATLN
EDDSNTIITNYOIGSPDLMECIMDLVEIVESPFPNOLRTQDOLVRSQILRKHIG
GLAFLITNTQETKRVADYVEORIEAFRSMELVQMSDTEKNIKETILNGKLG
DTSIDREVLKRWSEIKYKVEFPRRIEIOIMLSHLTKEDEVILNFINDNKNLRKLSVG
VGNHNTNSTRAQSGSLSDLDVDQDMANETISAKMTDKILFLFGESDDPS
SIKDIYAFKKDLVYFPLENTNPMLANKS"
complement(7844..8836)
/gene="CG11037"
/locus_tag="Dmel CG11037"
/old_locus_tag="CG11037"
/map="78A1-78A1"
/db_xref="FLYBASE:FBgn0037038"
complement(join(7844..8100,8177..8836))
/gene="CG11037"
/locus_tag="Dmel CG11037"
/old_locus_tag="CG11037"
/product="CG11037-RA"
/db_xref="FLYBASE:FBgn0037038"
complement(join(7880..8100,8177..8834))
/gene="CG11037"
/locus_tag="Dmel CG11037"
/old_locus_tag="CG11037"
/note="CG11037 gene product from transcript CG11037-RA"

```

```

/codon_start=1
/product="CG11037-PA"
/protein_id="AAF51662.2"
/db_xref="GI:23094197"
/db_xref="FLYBASE:FBgn0037038"
/translation="MEMHLVCCVLACTLVSTQVYGGDEKIVLSLPESQGRKNTLT

```

Query Match 98.2%; Score 5051.8; DB 2; Length 313157;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 5097; Conservative 0; Mismatches 17; Indels 5; Gaps 3;

```

QY 29 GTTCCCGCCAGAACGAGACGAAATAGACATTGCAACGATGGCTTATGGAAAGCCCAACGCA 88
Db 225546 GTTCCCGCCAGAACGAGACGAAATAGACATTGCAACGATGGCTTATGGAAAGCCCAACGCA 225605
QY 89 CCTGATGAGCTTTATCCAGCAAAATGCCCCAAGCGGAGCATGCGTGTGAAGGTACAGCTCGC 148
Db 225606 CCTGATGAGCTTTATCCAGCAAAATGCCCCAAGCGGAGCATGCGTGTGAAGGTACAGCTCGC 225665
QY 149 CGAAGATCTGGTGAACATTCCTTACGACGACACAACTCAATTTGTGCAACGACATGGG 208
Db 225666 CGAAGATCTGGTGAACATTCCTTACGACGACACAACTCAATTTGTGCAACGACATGGG 225725
QY 209 ATTCTTATTTGACGGTTTGATGCGATGGCTGACGGGCGACCACTTTAAGATTGCACAAAA 268
Db 225726 ATTCTTATTTGACGGTTTGATGCGATGGCTGACGGGCGACCACTTTAAGATTGCACAAAA 225785
QY 269 GTCCCTGAGAGCGTTCTCGAGCTAATTAAGCGATGGGCGAGCTTTTATGCAATACAC 328
Db 225786 GTCCCTGAGAGCGTTCTCGAGCTAATTAAGCGATGGGCGAGCTTTTATGCAATACAC 225845
QY 329 GGCCTACCGTTCTGCGACATGTGATCGATCGGCTGGGAGACAGCAGGAGCAACAGTCCGCA 388
Db 225846 GGCCTACCGTTCTGCGACATGTGATCGATCGGCTGGGAGACAGCAGGAGCAACAGTCCGCA 225905
QY 389 GAAGGCGCAACTTTGCTGCGCGACCTTCATGAGACACAGAGTCTTCCGCCCAAGGCGCT 448
Db 225906 GAAGGCGCAACTTTGCTGCGCGACCTTCATGAGACACAGAGTCTTCCGCCCAAGGCGCT 225965
QY 449 GATGCAACAGCTGGSCACTAGCTGCTTCAAGCACAAGACGCGGAGTGGCGGAGAGTT 508
Db 225966 GATGCAACAGCTGGSCACTAGCTGCTTCAAGCACAAGACGCGGAGTGGCGGAGAGTT 226025
QY 509 CCTTACAGCATTTGTGACGCTCTTCATGAGTACGGGACCCAGAGCTTAGTGTTCGCT 568
Db 226026 CCTTACAGCATTTGTGACGCTCTTCATGAGTACGGGACCCAGAGCTTAGTGTTCGCT 226085
QY 569 CTATATACCAACGATTTGTGACCTTTCGAGATCCACAGTTATGTGAGGAGGCGCGC 628
Db 226086 CTATATACCAACGATTTGTGACCTTTCGAGATCCACAGTTATGTGAGGAGGCGCGC 226145
QY 629 CATCCAAACGCTAGTGAATACTCAAGCATGTAGGGGATCGATTTGGGCCAGACCTCCG 688
Db 226146 CATCCAAACGCTAGTGAATACTCAAGCATGTAGGGGATCGATTTGGGCCAGACCTCCG 226205
QY 689 TCGCATGAGCATGTTCTGCTCGCTCGAATTGGCTATGTGAGGCAAAAGTTGCAACAGT 748
Db 226206 TCGCATGAGCATGTTCTGCTCGCTCGAATTGGCTATGTGAGGCAAAAGTTGCAACAGT 226265
QY 749 CAAACAGAGAGGTTACTGCTACTTCAAGCCCTTAAAAACAGATGGCAATGAGTGGG 808
Db 226266 CAAACAGAGAGGTTACTGCTACTTCAAGCCCTTAAAAACAGATGGCAATGAGTGGG 226325
QY 809 CTTGAGAGAGGCGCAATATTGGGTTGAGGAGAGCGACCAACGATGATTAAGCGGCGC 868
Db 226326 CTTGAGAGAGGCGCAATATTGGGTTGAGGAGAGCGACCAACGATGATTAAGCGGCGC 226385
QY 869 ACTACACTCGGCGGTTTGTGATCACTGCGGCCAAAACCAATGTGAACATGTGACCGG 928
Db 226386 ACTACACTCGGCGGTTTGTGATCACTGCGGCCAAAACCAATGTGAACATGTGACCGG 226445
QY 929 TGATGCGGCGGCGCTAACCATGGAATCTTTGGAATCTAGCTTTGAGGTGCTCCCGCAATT 988

```

```

Db 226446 TGATGCGGCGGCGCTAACCATGGAATCTTTGGAATCTAGCTTTGAGGTGCTCCCGCAATT 226505
QY 989 GAACATCTTCCACGCTTAAGAGACATGAGACGATATCTAACAGCAAGTACTAGTATCATCAG 1048
Db 226506 GAACATCTTCCACGCTTAAGAGACATGAGACGATATCTAACAGCAAGTACTAGTATCATCAG 226565
QY 1049 TGATTAACAAACGACATGAGGAGAAACGTGTGATGCTCTCCAGAAATCATGAGGCAATTGCT 1108
Db 226566 TGATTAACAAACGACATGAGGAGAAACGTGTGATGCTCTCCAGAAATCATGAGGCAATTGCT 226625
QY 1109 CATTCTACGCTATCAACCTCAGCGCGAGTTTGTGCTGTACAGCTTAAGAAATTTGCTGTT 1168
Db 226626 CATTCTACGCTATCAACCTCAGCGCGAGTTTGTGCTGTACAGCTTAAGAAATTTGCTGTT 226685
QY 1169 AAGCTTCTGAGACATCTCTAAGAGAGAACTACAGATCAAGTGTATCCGAGGCGTGCAAT 1228
Db 226686 AAGCTTCTGAGACATCTCTAAGAGAGAACTACAGATCAAGTGTATCCGAGGCGTGCAAT 226745
QY 1229 CACCATGCGCTCATATGTCTTAAGACGCTGAGAAATTAACCTAGATGCTTCTGCTGAGCAT 1288
Db 226746 CACCATGCGCTCATATGTCTTAAGACGCTGAGAAATTAACCTAGATGCTTCTGCTGAGCAT 226805
QY 1289 TTTGAGACACCTGATTAATTTAATACAGAACAGCGGAGAGTCAATTGCAATCCGCTTCCAC 1348
Db 226806 TTTGAGACACCTGATTAATTTAATACAGAACAGCGGAGAGTCAATTGCAATCCGCTTCCAC 226865
QY 1349 AATAGCTGTAAGTATATCATTAAGTATACATATGACACCAAGCTGCTCAAGATCTTACAC 1408
Db 226866 AATAGCTGTAAGTATATCATTAAGTATACATATGACACCAAGCTGCTCAAGATCTTACAC 226925
QY 1409 AGACATCTGTAATCAATCAAAATGCGAAGACATTAAGTCCACACTGTGTGAGTGTGCT 1468
Db 226926 AGACATCTGTAATCAATCAAAATGCGAAGACATTAAGTCCACACTGTGTGAGTGTGCT 226985
QY 1469 GCGTCTCTTGAAGAGTGGCAGACGAAAGCGTTGGAAGAAATGCCAGCTTAAGAGGA 1528
Db 226986 GCGTCTCTTGAAGAGTGGCAGACGAAAGCGTTGGAAGAAATGCCAGCTTAAGAGGA 227045
QY 1529 CACCTTAAAAAAATCCATTGGCGATGCGACATGCGATGCGACGCGCATTTCCAGATACGC 1588
Db 227046 CACCTTAAAAAAATCCATTGGCGATGCGACATGCGATGCGACGCGCATTTCCAGATACGC 227105
QY 1589 CTATTTGGGCTTTTCAGGCGTCACTTTCCAGAGCTGGCGGATCAAAATATATGAACATTAGA 1648
Db 227106 CTATTTGGGCTTTTCAGGCGTCACTTTCCAGAGCTGGCGGATCAAAATATATGAACATTAGA 227165
QY 1649 CATAGCTGCGCAGCGCGCATTTGAAGAGGAGCGGCGGAGAGAGGAACTGG 1708
Db 227166 CATAGCTGCGCAGCGCGCATTTGAAGAGGAGCGGCGGAGAGAGGAACTGG 227225
QY 1709 TACTGGGAGCTGGGAGCTGCACTGAAACGAGACGCACTGTATCCCGATTTGGCGGAACAC 1768
Db 227226 TACTGGGAGCTGGGAGCTGCACTGAAACGAGACGCACTGTATCCCGATTTGGCGGAACAC 227285
QY 1769 TGAACCTCTGCAAAAGCCCAACGCTAGTATGAGATCAATTTCAAGCGGTGACACTCGCGC 1828
Db 227286 TGAACCTCTGCAAAAGCCCAACGCTAGTATGAGATCAATTTCAAGCGGTGACACTCGCGC 227345
QY 1829 TGGTCAAGAGGCGCAAGTTAGAGCGGCAATATCACTATATTTCCAGGCAAGAAACCTTT 1888
Db 227346 TGGTCAAGAGGCGCAAGTTAGAGCGGCAATATCACTATATTTCCAGGCAAGAAACCTTT 227405
QY 1889 AGGACCTTAATTAATTTCCAACAGGACATGATGACAGAGGACAGGACAGTGTGATCACTGCGC 1948
Db 227406 AGGACCTTAATTAATTTCCAACAGGACATGATGACAGAGGACAGGACAGTGTGATCACTGCGC 227465
QY 1949 CAGACTCGCTGTAATTTCCAAATAGCGGTGCAACCACTTAACAAGCCGCGGATCGGTTAC 2008
Db 227466 CAGACTCGCTGTAATTTCCAAATAGCGGTGCAACCACTTAACAAGCCGCGGATCGGTTAC 227525
QY 2009 ACCAGGCGCGCGGGGAGCGGAGAGATCTAGTCAACGCGAGATCGAGATCCACCTTC 2068
Db 227526 ACCAGGCGCGCGGGGAGCGGAGAGATCTAGTCAACGCGAGATCGAGATCCACCTTC 227585

```


QY 2069 GCCAAGCAGCAAACTGAGGGATCACTACGGTGTATTGGTAATTATCTATTCGTGGGGCCAC 2128
 Db 227586 GCCAAGCAGCAAAACTGAGGGATCACTACGGTGTATTGGTAATTATCTATTCGTGGGGCCAC 227645
 QY 2129 TGGGGCCATACCCAAAAGGGCTCCGGAATACCCCGAAGCACACCGCTCTAGGAAAC 2188
 Db 227646 TGGGGCCATACCCAAAAGGGCTCCGGAATACCCCGAAGCACACCGCTCTAGGAAAC 227705
 QY 2189 GAGTCACACTAGTACGTGTGTGGCTTGTATGAAAACGAGTATGTACTCTACAGGTGCGG 2248
 Db 227706 GAGTCACACTAGTACGTGTGTGGCTTGTATGAAAACGAGTATGTACTCTACAGGTGCGG 227765
 QY 2249 GTTCGACGTAACGCCCGAGAGAAACAACCACTAAGACCAATGCGCGCGGACGACTGT 2308
 Db 227766 GTTCGACGTAACGCCCGAGAGAAACAACCACTAAGACCAATGCGCGCGGACGACTGT 227825
 QY 2309 GGGCCAAATCCCGTAAGAGAAACATACATTAGGGCTCGGAGATGATGACACGACGACTA 2368
 Db 227826 GGGCCAAATCCCGTAAGAGAAACATACATTAGGGCTCGGAGATGATGACACGACGACTA 227885
 QY 2369 TGTTCGCGGGACTACATGCGCAGCGGGGAAATGCGATGGGTAGAAAGCTCATGGAGC 2428
 Db 227886 TGTTCGCGGGACTACATGCGCAGCGGGGAAATGCGATGGGTAGAAAGCTCATGGAGC 227945
 QY 2429 TGATGAGTCTGTATGATCACTCCAGAGCCAGTTCTGTGTGTTCAAGAACGATCTTTCGA 2488
 Db 227946 TGATGAGTCTGTATGATCACTCCAGAGCCAGTTCTGTGTGTTCAAGAACGATCTTTCGA 228005
 QY 2489 CTCACGCTACCTAGAGTAATTAATGAACTACTCTATAGCGGGAGCCACACCGCTT 2548
 Db 228006 CTCACGCTACCTAGAGTAATTAATGAACTACTCTATAGCGGGAGCCACACCGCTT 228065
 QY 2549 GGACTGAGAGCAGCAGCGGGGCGCATTTGACGATGAGACGATTAATTCAGTTCTGCGC 2608
 Db 228066 GGACTGAGAGCAGCAGCGGGGCGCATTTGACGATGAGACGATTAATTCAGTTCTGCGC 228125
 QY 2609 ATGACGCAATGTTGTGTAAGAAAGATGGCTGTATGACCTCAACACAGTATCTGGCGA 2668
 Db 228126 ATGACGCAATGTTGTGTAAGAAAGATGGCTGTATGACCTCAACACAGTATCTGGCGA 228185
 QY 2669 TGGAAAGAGCTCACTGACGACGAACTTAATGGCTTGGACATGTTCCGCAAGATGT 2728
 Db 228186 TGGAAAGAGCTCACTGACGACGAACTTAATGGCTTGGACATGTTCCGCAAGATGT 228245
 QY 2729 CATGATATACACACCAAGGTGTACTGACTGTTTGGACACGCTCACCGAATGATCT 2788
 Db 228246 CATGATATACACACCAAGGTGTACTGACTGTTTGGACACGCTCACCGAATGATCT 228305
 QY 2789 GGTTCATGCGAAGCAAACTTCACGAAATGGCTCTTCATCTTTTGAACGCGGTTGTTCAA 2848
 Db 228306 GGTTCATGCGAAGCAAACTTCACGAAATGGCTCTTCATCTTTG-TTGACGCGGTTGTTCAA 228362
 QY 2849 CAATTTGGGACCGATCTACTCATTTGATGACACGCAAGATTTGGAAGCGCTACAGGT 2908
 Db 228363 CAATTTGGGACCGATCTACTCATTTGATGACACGCAAGATTTGGAAGCGCTACAGGT 228422
 QY 2909 GGTCACGAATATTTCCCAAGCAGTTGACAGCTTAAGAACTATTTGAATCATATGGA 2968
 Db 228423 GGTCACGAATATTTCCCAAGCAGTTGACAGCTTAAGAACTATTTGAATCATATGGA 228482
 QY 2969 TTCTACTCAAAACGCAACTACCAAGACGCGATGCGCATTCGCGCTTCTTACGGAATCT 3028
 Db 228483 TTCTACTCAAAACGCAACTACCAAGACGCGATGCGCATTCGCGCTTCTTACGGAATCT 228542
 QY 3029 GGCCAAATACATCTGTAAAAGCAGCACTTTCCAGACGACCAAGGCGCTGCGAGCG 3088
 Db 228543 GGCCAAATACATCTGTAAAAGCAGCACTTTCCAGACGACCAAGGCGCTGCGAGCG 228602
 QY 3089 GACGGTCTTAAAGCTGAGCCAGCTGCGCGGAGATCAGAACTGATGAGCTGCGCTCCA 3148
 Db 228603 GACGGTCTTAAAGCTGAGCCAGCTGCGCGGAGATCAGAACTGATGAGCTGCGCTCCA 228662

QY 3149 GGCAGAGCTGCTTGTGTGAGCTGCTGTATTAACCTGAATACCCGCAATGACCTTTTACT 3208
 Db 228663 GGCAGAGCTGCTTGTGTGAGCTGCTGTATTAACCTGAATACCCGCAATGACCTTTTACT 228722
 QY 3209 GGCAGACTGCGCAAAAGTATATCAGAGACTGCGCCGATCCTGTGATCATTCGACATGAG 3268
 Db 228723 GGCAGACTGCGCAAAAGTATATCAGAGACTGCGCCGATCCTGTGATCATTCGACATGAG 228782
 QY 3269 GCGGCAAAACCAAGTTGCAATTCGGGTGCAATTCGCTAGTACCTTCATTTAGAGCA 3328
 Db 228783 GCGGCAAAACCAAGTTGCAATTCGGGTGCAATTCGCTAGTACCTTCATTTAGAGCA 228842
 QY 3329 TAGCACTCCAAAGCTTTTGCAAGTCCCTGTGTGGGTCCATTTGGCTTCGCTTACAGCCA 3388
 Db 228843 TAGCACTCCAAAGCTTTTGCAAGTCCCTGTGTGGGTCCATTTGGCTTCGCTTACAGCCA 228902
 QY 3389 CCACGACCAACTTAGATCAGCTCTACTGTCAGAGCTCCCGGCACTTTCGCTGAGGCA 3448
 Db 228903 CCACGACCAACTTAGATCAGCTCTACTGTCAGAGCTCCCGGCACTTTCGCTGAGGCA 228962
 QY 3449 GGAGCTGCTCTTTTCTCGAGCTGACATTCAGCAACAATTACAGAGACGTGAGGA 3508
 Db 228963 GGAGCTGCTCTTTTCTCGAGCTGACATTCAGCAACAATTACAGAGACGTGAGGA 229022
 QY 3509 GATCCGGCACTGCTTGGCGGTCAAGTACCAAGACGCGCTGGCGCCAAATGGCTTCAATGG 3568
 Db 229023 GATCCGGCACTGCTTGGCGGTCAAGTACCAAGACGCGCTGGCGCCAAATGGCTTCAATGG 229082
 QY 3569 ACACTTGAGTATACAGATCAAGGAGCAAGAGATTCGTGTGATCCCTGCTTCCAACTC 3628
 Db 229083 ACACTTGAGTATACAGATCAAGGAGCAAGAGATTCGTGTGATCCCTGCTTCCAACTC 229142
 QY 3629 CAAGAGCAATGTCGCGCCCAACATCAACCACTCAATACCTGAGTACAGCAACAATGAG 3688
 Db 229143 CAAGAGCAATGTCGCGCCCAACATCAACCACTCAATACCTGAGTACAGCAACAATGAG 229202
 QY 3689 GCTGATTAATCTGAGACGGGAAAGAACCTCAGAAACGCAAGTACCACTGACGATGC 3748
 Db 229203 GCTGATTAATCTGAGACGGGAAAGAACCTCAGAAACGCAAGTACCACTGACGATGC 229262
 QY 3749 CAAGGATCAAGGATCGATTAATATGAGCTGAGAAATGAGAGCTGATCTGAGCCAGCA 3808
 Db 229263 CAAGGATCAAGGATCGATTAATATGAGCTGAGAAATGAGAGCTGATCTGAGCCAGCA 229322
 QY 3809 CCTGATGAGAGCGAAGTGTGCTGTGAGCTTGAACGCTAACAAAGATCAACCGCTGA 3868
 Db 229323 CCTGATGAGAGCGAAGTGTGCTGTGAGCTTGAACGCTAACAAAGATCAACCGCTGA 229382
 QY 3869 GTTGCTTCAAGCTCACTTAACCTGGGATTTTGATCAAGGGCGGAAACTGTGAGCT 3928
 Db 229383 GTTGCTTCAAGCTCACTTAACCTGGGATTTTGATCAAGGGCGGAAACTGTGAGCT 229442
 QY 3929 GCCCAATAGCACTTAACTGATCATGAGGATGCGCTTAACATTCGAGGCGGAGCA 3988
 Db 229443 GCCCAATAGCACTTAACTGATCATGAGGATGCGCTTAACATTCGAGGCGGAGCA 229502
 QY 3989 TAGGAGCTGTATCGTGTGCGCTGACAGCTGCTAGTAAGATTAGAGGACAACTAAT 4048
 Db 229503 TAGGAGCTGTATCGTGTGCGCTGACAGCTGCTAGTAAGATTAGAGGACAACTAAT 229562
 QY 4049 GGGTCACAACTGATGACATTTCTAGAGCTGATTTTGTGGAAGATCATCAAGTGCATTA 4108
 Db 229563 GGGTCACAACTGATGACATTTCTAGAGCTGATTTTGTGGAAGATCATCAAGTGCATTA 229622
 QY 4109 AACAGCAAGAGAGGCTTTGGCGGATATGCACTGATGATACCAAGATATGACCAATCCTT 4168
 Db 229623 AACAGCAAGAGAGGCTTTGGCGGATATGCACTGATGATACCAAGATATGACCAATCCTT 229682
 QY 4169 GCGCTGATCTGTCACTCAACATTTGATCCGATTTTGGAGCGGATGAGTTTCCAC 4228
 Db 229683 GCGCTGATCTGTCACTCAACATTTGATCCGATTTTGGAGCGGATGAGTTTCCAC 229742
 QY 4229 AATTTTGTGCGGCATCAAGATCTGTGGAAGTACCAACCACTGTTGGAATATC 4288

DB 229743 AAATTTGTCGGCATCAAGATCTGCTGGAGGACCCGAACACAGATGTTGGAGTTC 229802
 QY 4289 AGACGGCCCACTTGAGCATTTGTTCCCATCATCGGCGCATCGGCGAGACACGCAATC 4348
 DB 229803 AGAGCCCACTTGAGCATTTGTTCCCATCATCGGCGCATCGGCGAGACACGCAATC 229862
 QY 4349 GATGTTCCCAAGCTGCGGTCTTCTGCAATCGTCAAGCTGTACTTTGTTCTGGCGCAGGA 4408
 DB 229863 GATGTTCCCAAGCGCGGTCTTCTGCAATCGTCAAGCTGTACTTTGTTCTGGCGCAGGA 229922
 QY 4409 GAAAGTCACCGCAAGCTGTCACTGCTTAATCCACCAAGTTAAGCTCTTCAAGCTGT 4468
 DB 229923 GAAGGTCAAGCCCAAGCTGTCACTGCTTAATCCACCAAGTTAAGCTCTTCAAGCTGT 229982
 QY 4469 CATCGAAGACCGGAGCATCGATCGTGGGGAGAGAGCTCTACCAAGAACTCTCTCCG 4528
 DB 229983 CATCGAAGACCGGAGCATCGATCGTGGGGAGAGAGCTCTACCAAGAACTCTCTCCG 230042
 QY 4529 GGCATGTCGTCAATGATGCGGAGCCCTTAATAGATTCCTGCTC--TCGTCGCCACCA 4586
 DB 230043 GGCATGTCGTCAATGATGCGGAGCCCTTAATAGATTCCTGCTCAGACACACCA 230102
 QY 4587 ACAAGACACAGACGCGGTGCTCTCTCGGCTGAGAGAGAGTGAAGCGCGGACAT 4646
 DB 230103 ACAAGACACAGACGCGGTGCTCTCTCGGCTGAGAGAGAGTGAAGCGCGGACAT 230162
 QY 4647 TAAAT 4706
 DB 230163 TAAAT 230222
 QY 4707 ACTTGTGTAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4766
 DB 230223 ACTTGTGTAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 230282
 QY 4767 AAGGCGCATGCGCCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4826
 DB 230283 AAGGCGCATGCGCCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 230342
 QY 4827 GCGAGTAATTCGCGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4886
 DB 230343 GCGAGTAATTCGCGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 230402
 QY 4887 ATCAAGTGTAGCTCTTTTACAGAGCTCTTATATATATATATATATATATATATATATAT 4946
 DB 230403 ATCAAGTGTAGCTCTTTTACAGAGCTCTTATATATATATATATATATATATATATAT 230462
 QY 4947 AGCTTATATACGCTCTTATGAGAACTGACAGCTGCTTCTTCTTCTTCTTCTTCTTCA 5006
 DB 230463 AGCTTATATACGCTCTTATGAGAACTGACAGCTGCTTCTTCTTCTTCTTCTTCTTCA 230522
 QY 5007 GACTTGTATCTTATATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 5066
 DB 230523 GACTTGTATCTTATATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 230582
 QY 5067 TTCTGCTCTTACAGCTAT 5126
 DB 230583 TTCTGCTCTTACAGCTAT 230642
 QY 5127 GATTAAATCAATAAAAA 5145
 DB 230643 GATTAAATCAATAAAAA 230661

ORGANISM

Rattus norvegicus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 96091)
 Muzny D, Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alibekov, S., Amin, A., Angiano, D., Anyledechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benhammed, F., Bismato, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Ande, C., Dedetich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinu, L., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durkin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacod, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Koyak, K., Kowicz, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorusubewa, L., Louisedge, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. E., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munda, M., Murphy, M., Nat, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackelemb, O., Okomou, G., Olarinmagaon, A., Pal, S., Parks, K., Paeternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Pu, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, A., Reish, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Rhodes, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svetek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umanak, K., Vais, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Woodem, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zuo, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE

JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE

Rat Genome Sequencing Consortium.
 Submitted (10-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 96091)

COMMENT

On Oct 10, 2002 this sequence version replaced gi:2218203.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas

Db	44550	GGCTACCGTTCTGCCAATGTGATGATGGCTGGGAGACAGCGGACACAGTCCGCA	44491
QY	389	GAAGCGCACTTCTGTGCGGACCTCATGAGCAAGAGTCTCCGCCAGGCGCT	448
Db	44490	GAAGCGCAACTTCTGTGCGGACCTCATGAGCAAGAGTCTCCGCCAGGCGCT	44431
QY	449	GATGACAAGCTGGCCACTAGTGGTTCAACACAAGCAAGCAAGTGGCGCAGAGATT	508
Db	44430	GATGACAAGCTGGCCACTAGTGGTTCAACACAAGCAAGCAAGTGGCGCAGAGATT	44371
QY	509	CCTTGACAGATTGTGAACCGCTTTCATGAGTAAGGACCCAGACGTTAGTGGCGT	568
Db	44370	CCTTGACAGATTGTGAACCGCTTTCATGAGTAAGGACCCAGACGTTAGTGGCGT	44311
QY	569	CTATTATCACCAAGTTTGTGCACTTCTCGAGATCCACAGTTAATGTAGGAGGCGCG	628
Db	44310	CTATTATCACCAAGTTTGTGCACTTCTCGAGATCCACAGTTAATGTAGGAGGCGCG	44251
QY	629	CATCCAAACGCTAGTGAATCTCAAGACATGTAGGGGATTCGATTGGCCACACTCCG	688
Db	44250	CATCCAAACGCTAGTGAATCTCAAGACATGTAGGGGATTCGATTGGCCACACTCCG	44191
QY	689	TCGCATGACGATGTTCCGTCCCTCGAATTGGCTATATTGGACAAAGTTGGACCGGT	748
Db	44190	TCGCATGACGATGTTCCGTCCCTCGAATTGGCTATATTGGACAAAGTTGGACCGGT	44131
QY	749	CAAAACGAGGGGTCTACTGCTACCTTCAGCCCTTAAACACGACATGGCAATGAGTGG	808
Db	44130	CAAAACGAGGGGTCTACTGCTACCTTCAGCCCTTAAACACGACATGGCAATGAGTGG	44071
QY	809	CTTGGACGAGGCCGCACAATATTTGGGTTGAGGAGCCACCAACGATGATTAGCGGCC	868
Db	44070	CTTGGACGAGGCCGCACAATATTTGGGTTGAGGAGCCACCAACGATGATTAGCGGCC	44011
QY	869	ACTACACTCGGCGCTTTCGTCACTACTGCGCCCAAAACCAATGTAAAGATGTACCGG	928
Db	44010	ACTACACTCGGCGCTTTCGTCACTACTGCGCCCAAAACCAATGTAAAGATGTACCGG	43951
QY	929	TGATGCCGGCGCGTAAACATGNAATCTTTCGATCTAGCTTTGAGGTGGTCCGCAATT	988
Db	43950	TGATGCCGGCGCGTAAACATGNAATCTTTCGATCTAGCTTTGAGGTGGTCCGCAANT	43891
QY	989	GAACATCTTCCACGCTTAAGACATGACGATATCTCAAGCAAGTACTAGATCATCAG	1048
Db	43890	GAACATCTTCCACGCTTAAGACATGACGATATCTCAAGCAAGTACTAGATCATCAG	43831
QY	1049	TGATTAACACGACAGCTGGAGGAAACGTGTGATGCTCTTCAGAAATTAAGGGCATTGGT	1108
Db	43830	TGATTAACACGACAGCTGGAGGAAACGTGTGATGCTCTTCAGAAATTAAGGGCATTGGT	43771
QY	1109	CATTCTCAGCTATCACACTCAGCGCGAGTTTGTGCTGTACAGCTTAAAGAAATTGGCT	1168
Db	43770	CATTCTCAGCTATCACACTCAGCGCGAG-TTGTGCTGTACAGCTTAAAGAAATTGGCT	43712
QY	1169	AACTTGTGAGCACTCTCAAGAGGAATCAAGATCAAGGTATCCGAGGCGTGCAT	1228
Db	43711	AACTTGTGAGCACTCTCAAGAGGAATCAAGATCAAGGTATCCGAGGCGTGCAT	43652
QY	1229	CACCATGCGCTACATGCTCTAAGACGCTGGAATAATCAATGCTTCTGTGAGCAT	1288
Db	43651	CACCATGCGCTACATGCTCTAAGACGCTGGAATAATCAATGCTTCTGTGAGCAT	43592
QY	1289	TTTGGACACCTGTTAATTTAATACAGAAACAGCGGAGTGCATTGCACTCCGCTTCAC	1348
Db	43591	TTTGGACACCTGTTAATTTAATACAGAAACAGCGGAGTGCATTGCACTCCGCTTCAC	43532
QY	1349	AATAGCTCTGAAGTATATCTTAAGTATACATGACCAAAAGTCTTCAAGATCTACAC	1408
Db	43531	AATAGCTCTGAAGTATATCTTAAGTATACATGACCAAAAGTCTTCAAGATCTACAC	43472
QY	1409	AGACACTCTGAATCAATCAAGTCAGAGGACATTAAGTCCACACTGTGTGAGCTATGGT	1468

[illegible]

TITLE
JOURNAL
COMMENT

Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Kearney, L., Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M.A., Madada, P., Moshrefi, A.R., Moshrefi, M., Nixon, K., Paclebo, J.M., Park, S., Pfeiffer, B., Poon, L., Segueira, A., Sethi, H., Snir, E., Svirskaas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and Rudin, G.M.

Direct Submission
Submitted (18-FEB-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-111, Berkeley, CA 94720, USA
On Aug 2, 1999 this sequence version replaced gi:5629992.

For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to bdg@fruitfly.berkeley.edu. All contigs in this submission meet the following cutoffs: length >= 200 bases.

* NOTE: This is a 'working draft' sequence. It currently consists of 82 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1
571 570: contig of 570 bp in length
651 650: gap of unknown length
1422 1421: contig of 771 bp in length
1502 1501: gap of unknown length
2631 2630: contig of 1129 bp in length
2710 2710: gap of unknown length
3461 3461: contig of 751 bp in length
3542 3542: gap of unknown length
4249 4249: contig of 708 bp in length
4329 4329: gap of unknown length
4330 4330: contig of 1076 bp in length
5405 5405: gap of unknown length
5485 5485: gap of unknown length
5486 5486: gap of 645 bp in length
6130 6130: contig of 754 bp in length
6210 6210: gap of unknown length
6211 6211: gap of 754 bp in length
6964 6964: gap of unknown length
7044 7044: gap of 816 bp in length
7045 7045: gap of unknown length
7861 7861: gap of 971 bp in length
7941 7941: gap of unknown length
8911 8911: gap of unknown length
8912 8912: gap of unknown length
8992 8992: contig of 1380 bp in length
10372 10372: gap of unknown length
10452 10452: contig of 720 bp in length
11171 11171: gap of unknown length
11251 11251: gap of 189 bp in length
1172 1172: gap of unknown length
12640 12640: contig of 815 bp in length
12720 12720: gap of unknown length
12721 12721: gap of 815 bp in length
13535 13535: gap of unknown length
13615 13615: gap of 631 bp in length
14246 14246: gap of unknown length
14326 14326: gap of 932 bp in length
14327 14327: gap of unknown length
15258 15258: gap of 699 bp in length
15338 15338: gap of unknown length
16037 16037: gap of 646 bp in length
16117 16117: gap of unknown length
16763 16763: gap of 777 bp in length
16843 16843: gap of unknown length
17620 17620: gap of 814 bp in length
17700 17700: gap of unknown length
18514 18514: gap of 682 bp in length
18594 18594: gap of unknown length
19276 19276: gap of 888 bp in length
19356 19356: gap of unknown length
20244 20244: gap of 567 bp in length
20324 20324: gap of unknown length
20891 20891: gap of 674 bp in length
20972 20972: gap of unknown length
21645 21645: gap of 956 bp in length
21725 21725: gap of unknown length
22681 22681: gap of 728 bp in length
22761 22761: gap of unknown length

22762 22762: contig of 728 bp in length
23490 23490: gap of unknown length
23570 23570: contig of 866 bp in length
24435 24435: gap of unknown length
24515 24515: gap of 1325 bp in length
25840 25840: gap of 1138 bp in length
25920 25920: gap of unknown length
25921 25921: gap of unknown length
27058 27058: gap of 1018 bp in length
27138 27138: gap of 1018 bp in length
27159 27159: gap of 948 bp in length
28157 28157: gap of unknown length
28237 28237: gap of 897 bp in length
29185 29185: gap of unknown length
29265 29265: gap of 1048 bp in length
30162 30162: gap of unknown length
30242 30242: gap of 1035 bp in length
31290 31290: gap of 1474 bp in length
32405 32405: gap of unknown length
32485 32485: gap of 1329 bp in length
33959 33959: gap of unknown length
34039 34039: gap of 1757 bp in length
35447 35447: gap of unknown length
35468 35468: gap of 1882 bp in length
37205 37205: gap of unknown length
37285 37285: gap of 1894 bp in length
39167 39167: gap of unknown length
39247 39247: gap of 1818 bp in length
41141 41141: gap of unknown length
41221 41221: gap of 1580 bp in length
43039 43039: gap of unknown length
43119 43119: gap of 2760 bp in length
44699 44699: gap of unknown length
44778 44778: gap of 3418 bp in length
47538 47538: gap of unknown length
47539 47539: gap of 617 bp in length
47619 47619: gap of unknown length
51036 51036: gap of 621 bp in length
51037 51037: gap of 624 bp in length
51117 51117: gap of 624 bp in length
51734 51734: gap of 624 bp in length
51814 51814: gap of 624 bp in length
52435 52435: gap of 624 bp in length
52515 52515: gap of 624 bp in length
53201 53201: gap of 624 bp in length
53281 53281: gap of 624 bp in length
53750 53750: gap of 624 bp in length
53830 53830: gap of 624 bp in length
54276 54276: gap of 624 bp in length
54356 54356: gap of 624 bp in length
54979 54979: gap of 624 bp in length
54980 54980: gap of 624 bp in length
55059 55059: gap of 624 bp in length
55681 55681: gap of 624 bp in length
55761 55761: gap of 624 bp in length
55762 55762: gap of 624 bp in length
56385 56385: gap of 624 bp in length
56465 56465: gap of 624 bp in length
57182 57182: gap of 624 bp in length
57262 57262: gap of 624 bp in length
57812 57812: gap of 624 bp in length
57863 57863: gap of 624 bp in length
57892 57892: gap of 624 bp in length
58558 58558: gap of 624 bp in length
58638 58638: gap of 624 bp in length
58659 58659: gap of 624 bp in length
59362 59362: gap of 624 bp in length
59363 59363: gap of 624 bp in length
59443 59443: gap of 624 bp in length
60109 60109: gap of 624 bp in length
60189 60189: gap of 624 bp in length
60813 60813: gap of 624 bp in length
60893 60893: gap of 624 bp in length
60894 60894: gap of 624 bp in length
61473 61473: gap of 624 bp in length
61553 61553: gap of 624 bp in length
61554 61554: gap of 624 bp in length
61554 61554: gap of 624 bp in length
62190 62190: gap of 624 bp in length
62270 62270: gap of 624 bp in length
62271 62271: gap of 624 bp in length
63085 63085: gap of 624 bp in length
63165 63165: gap of 624 bp in length
63166 63166: gap of 624 bp in length
63789 63789: gap of 624 bp in length
63790 63790: gap of 624 bp in length
63870 63870: gap of 624 bp in length
64565 64565: gap of 624 bp in length

64646 64645: gap of unknown length
* 64646 65257: contig of 612 bp in length
* 65337 65337: gap of unknown length
* 65338 65942: contig of 605 bp in length
* 65943 66022: gap of unknown length
* 66023 66659: contig of 637 bp in length
* 66660 66739: gap of unknown length
* 66740 67583: contig of 844 bp in length
* 67584 67663: gap of unknown length
* 67664 68269: contig of 606 bp in length
* 68270 68349: gap of unknown length
* 68350 69001: contig of 652 bp in length
* 69002 69081: gap of unknown length
* 69082 69984: contig of 883 bp in length
* 69985 70044: gap of unknown length
* 70045 70796: contig of 752 bp in length
* 70797 70876: gap of unknown length
* 70877 71572: contig of 696 bp in length
* 71573 71652: gap of unknown length
* 71653 72306: contig of 654 bp in length
* 72307 72386: gap of unknown length
* 72387 73045: contig of 659 bp in length
* 73046 73125: gap of unknown length
* 73126 73949: contig of 824 bp in length
* 73950 74029: gap of unknown length
* 74030 74595: contig of 566 bp in length

Query Match 35.7%; Score 1834.4; DB 14; Length 79733;
Best Local Similarity 95.2%; Pred. No. 0;

Matches 1864; Conservative 0; Mismatches 91; Indels 3; Gaps 2;

QY 1449 ACACGTGTGAGTGTGATGCTGCTCTTTCGAGAGTGTGCAAGAGCGTTGAAAG 1508
DB 41119 ACACGTGTGAGTGTGATGCTGCTCTTTCGAGAGTGTGCAAGAGCGTTGAAAG 41060
QY 1509 AATGCCACGCTACTAAGGAGACCTTAAATAATCCATTGGCGATGAGACCTGCGATGA 1568
DB 41059 AATGCCACGCTACTAAGGAGACCTTAAATAATCCATTGGCGATGAGACCTGCGATGA 41000
QY 1569 CGCGGCATTCAGATACGCTATTGGGCTTTTCAGGCGCTCACTTCCAGAGCTGGCGAT 1628
DB 40999 CGCGGCATTCAGATACGCTATTGGGCTTTTCAGGCGCTCACTTCCAGAGCTGGCGAT 40940
QY 1629 CAAATATATGGAACATTAGACATAGCTGCCACGCGCATTTAGAAAGGAAAGGAGG 1688
DB 40939 CAAATATATGGAACATTAGACATAGCTGCCACGCGCATTTAGAAAGGAAAGGAGG 40880
QY 1689 GCGGAGAGAGAGAACTGTACTGTGAGTGTGAGTGTGACCTGAACAGAGCACTGTA 1748
DB 40879 GCGGAGAGAGAGAACTGTACTGTGAGTGTGAGTGTGACCTGAACAGAGCACTGTA 40820
QY 1749 TCCCGCATTTGGCGGAGACCTGTGACCTGCAAAAGCCAGCGCTGATAGATCCATT 1808
DB 40819 TCCCGCATTTGGCGGAGACCTGTGACCTGCAAAAGCCAGCGCTGATAGATCCATT 40760
QY 1809 TCAGCGGTGAGACCTGCGGTGCTCAACAGAGCCAAAGTTAGAGCGGATATACATATAT 1868
DB 40759 TCAGCGGTGAGACCTGCGGTGCTCAACAGAGCCAAAGTTAGAGCGGATATACATATAT 40700
QY 1869 TCCAGCAAAAGAAACCTTTAGAGCTATATTTCCAAACAGGCAATGAGAGAGCA 1928
DB 40699 TCCAGCAAAAGAAACCTTTAGAGCTATATTTCCAAACAGGCAATGAGAGAGCA 40640
QY 1929 GAGGAGTGTGATCATCTGCCAGACCTGCGCTGAATTCCTCAATAGCGGTGACACAGCT 1988
DB 40639 GAGGAGTGTGATCATCTGCCAGACCTGCGCTGAATTCCTCAATAGCGGTGACACAGCT 40580
QY 1989 ACAACGCGGAGTGTGATCATCAACAGCGCCCGGAGAGCGGAGAGATATCTCAGTCAAG 2048
DB 40579 ACAACGCGGAGTGTGATCATCAACAGCGCCCGGAGAGCGGAGAGATATCTCAGTCAAG 40520
QY 2049 CAGAGATGAGATCCACTGCGCAACAGCAAACTAGAGATGATGATGATTTGGT 2108

DB 40519 CAGAGATGAGATCCACTGCGCAACAGCAAACTAGAGATGATGATGATTTGGT 40460
QY 2109 AATTACTATGTGGGAGCACTGGCGCCATACCCAAAAAGGCTCCGGAATATCCCGAAGC 2168
DB 40459 AATTACTATGTGGGAGCACTGGCGCCATACCCAAAAAGGCTCCGGAATATCCCGAAGC 40400
QY 2169 ACAGCCAGCTCTTAGGAAACGAGTCCAACTAGTCAAGTGTGCTTGAATGAACGAGT 2228
DB 40399 ACAGCCAGCTCTTAGGAAACGAGTCCAACTAGTCAAGTGTGCTTGAATGAACGAGT 40340
QY 2229 ATATCTCTTACAGGAGTGGGGGTCTGACGATAGCGCCGAGAGAAACACAGTAAAGCA 2288
DB 40339 ATATCTCTTACAGGAGTGGGGGTCTGACGATAGCGCCGAGAGAAACACAGTAAAGCA 40280
QY 2289 TCGGCGCGGAGCAAGTGTGCGCAATCCCGTGAAGCAATACATTAGGCGTGGGA 2348
DB 40279 TCGGCGCGGAGCAAGTGTGCGCAATCCCGTGAAGCAATACATTAGGCGTGGGA 40220
QY 2349 GATGATGAGACGACGATATGTTTCCGAGGACTACATGCGAGCGGAGATGCGATG 2408
DB 40219 GATGATGAGACGACGATATGTTTCCGAGGACTACATGCGAGCGGAGATGCGATG 40160
QY 2409 GGTAGGAGCTCATGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 2468
DB 40159 GGTAGGAGCTCATGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 40100
QY 2469 TGTTCAGAAAGATCTTTCAGCTCCAGCTACCTAGAGTAAATGAACTACCTCACTT 2528
DB 40099 TGTTCAGAAAGATCTTTCAGCTCCAGCTACCTAGAGTAAATGAACTACCTCACTT 40040
QY 2529 AGCGGAGGACCAACCGCTGTGAGTGTGAGAGACGAGCGGCGCATTTGAGCAATCGAG 2588
DB 40039 AGCGGAGGACCAACCGCTGTGAGTGTGAGAGACGAGCGGCGCATTTGAGCAATCGAG 39980
QY 2589 AGGATATTCAGTTCGCGCATGACGCAATGCTGGAAGGAAGATGGCGCTGATCAGC 2648
DB 39979 AGGATATTCAGTTCGCGCATGACGCAATGCTGGAAGGAAGATGGCGCTGATCAGC 39920
QY 2649 CTCACAGATATCTGGCCGATGGAAGAGCTCACTCAGCAGCACTTAAATGCTTCTG 2708
DB 39919 CTCACAGATATCTGGCCGATGGAAGAGCTCACTCAGCAGCACTTAAATGCTTCTG 39860
QY 2709 GACATGTTCCGAGAGATGTTCAATGATATACACACACAGAGTGTACTCAGTCTTTGGAC 2768
DB 39859 GACATGTTCCGAGAGATGTTCAATGATATACACACACAGAGTGTACTCAGTCTTTGGAC 39800
QY 2769 AGGTTCACCGAATGATTTGTTGTTCAATGGAAGAACTTCAAGAAATGGCTTCTATCT 2828
DB 39799 AGGTTCACCGAATGATTTGTTGTTCAATGGAAGAACTTCAAGAAATGGCTTCTATCT 39742
QY 2829 TGTTCAGCGGTTGTTCAAAATTTGGGACCGATCTCACTCAATGATGACAGCAAG 2888
DB 39741 TGTTCAGCGGTTGTTCAAAATTTGGGACCGATCTCACTCAATGATGACAGCAAG 39683
QY 2889 AATTGGAAGACGCTACAGTGTGTCAGAAATTTTCCGACGAGTTGAGCTTAAAGAA 2948
DB 39682 AATTGGAAGACGCTACAGTGTGTCAGAAATTTTCCGACGAGTTGAGCTTAAAGAA 39623
QY 2949 CTAATTAAGATATATGAGATTTCTACTCAAAAGCCAACTAACAGAGCGGATCGCATTT 3008
DB 39622 CTAATTAAGATATATGAGATTTCTACTCAAAAGCCAACTAACAGAGCGGATCGCATTT 39563
QY 3009 CTGCGCTTCCTTACGAGATCTGGCCATATCATCTGTTAAAGAGAGGAGCTTCCAGCGAG 3068
DB 39562 CTGCGCTTCCTTACGAGATCTGGCCATATCATCTGTTAAAGAGAGGAGCTTCCAGCGAG 39503
QY 3069 CAGAGCAGGCTGTGAGCGAGCGGTCTTAAAGCTGCGCCAGCTGCGCGGATCAAGAG 3128
DB 39502 CAGAGCAGGCTGTGAGCGAGCGGTCTTAAAGCTGCGCCAGCTGCGCGGATCAAGAG 39443
QY 3129 TCGATGAGAGTGTGCTCCAGGCGAGAGCTGCGCTGATGATGATGATGATGATGATG 3188
DB 39442 TCGATGAGAGTGTGCTCCAGGCGAGAGCTGCGCTGATGATGATGATGATGATGATG 39383

QY 3189 CCGAATATGACCTTTTACTGCGCCGACCTGCGAAGATATATCAGACTCTGCCGATCC 3248
DB 39382 CCGAATATGACCTTTTACTGCGCCGACCTGCGAAGATATATCAGACTCTGCCGATCC 39323
QY 3249 TGCATTCATTTGGACATAGGCGGCGCAAGCCAAAGTTGCAATTCGGGTGCCAATTCGGCT 3308
DB 39322 TGCATTCATTTGGACATAGGCGGCGCAAGCCAAAGTTGCAATTCGGGTGCCAATTCGGCT 39263
QY 3309 AGTAGCTCTGCATGAGCAGTAGAGCTCCCAAGCTTTGCAAGATCCCTCTGTGGGTCCA 3368
DB 39262 AGTAGCTCTGCATGAGCAGTAGAGCTCCCAAGCTTTGCAAGATCCCTCTGTGGGTCCA 39203
QY 3369 TTTGCTCTGCTTCAGAGCCACCACCACCACTTAGCAT 3406
DB 39202 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTT 39165

RESULT 12
AR495636 767 bp DNA linear PAT 22-SEP-2004
LOCUS Sequence 596 from patent US 6703491.
DEFINITION AR495636
ACCESSION AR495636
VERSION AR495636.1 GI:52431111
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 767)
AUTHORS Homburger, S.A., Ebens, A.J., Jr., Erickson, C.S., Francis-Lang, H.L.,
Margolis, J.S., Reddy, B.P., Ruddy, D.A. and Buchman, A.R.
TITLE Drosophila sequences
JOURNAL Patent: US 6703491-A 596 09-MAR-2004;
Exelixis, Inc.; South San Francisco, CA

FEATURES
source location/Qualifiers
1..767
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 14.2%; Score 732.8; DB 6; Length 767;
Best Local Similarity 98.7%; Pred. No. 9.1e-175;
Matches 760; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

QY 2639 CCTGATCAGCTTCACACAGTATCTGGCCGATGAAAGAGCTCACTAGCAGCAACTTAA 2698
DB 1 CCTGATCAGCTTCACACAGTATCTGGCCGATGAAAGAGCTCACTAGCAGCAACTTAA 60
QY 2699 ATGCGTTCTGGACATGTTCCGCAAGATGTTCAATGATACACACACCAAGTGTACTACT 2758
DB 61 ATGCGTTCTGGACATGTTCCGCAAGATGTTCAATGATACACACACCAAGTGTACTACT 120
QY 2759 GTTTTGGACACGGTACCGAAGTATCTGGTTCATGCGAAGAAATTCACGAATGG 2818
DB 121 GTTTTGGACACGGTACCGAAGTATCTGGTTCATGCGAAGAAATTCACGAATGG 178
QY 2819 CTCTTCATCTTGTGACGCGGTGTTTCACAAATTTGGGACCGATCTACTCAATTGGAT 2878
DB 179 CTCTTCATCTTGTGACGCGGTGTTTCACAAATTTGGGACCGATCTACTCAATTGGAT 237
QY 2879 GCACAGCAAGTTTGAAGACGCTACAGGTGCTCAGCAATATTTCCCGACGCGATTGCA 2938
DB 238 GCACAGCAAGTTTGAAGACGCTACAGGTGCTCAGCAATATTTCCCGACGCGATTGCA 297
QY 2939 GCTTAAAGAACTATTGAATCATATGGAATTTCTACTCAAAAGCCCACTACAGAGCGG 2998
DB 298 GCTTAAAGAACTATTGAATCATATGGAATTTCTACTCAAAAGCCCACTACAGAGCGG 357
QY 2999 CATGCCCATTTCTGGCTCTCTTAAGGATCTGGCCCAATCATATCTAATAAGACAGGACTT 3058
DB 358 CATGCCCATTTCTGGCTCTCTTAAGGATCTGGCCCAATCATATCTAATAAGACAGGACTT 417
QY 3059 TCCAGCGACCAAGACCGGCTTGAGCGGAGCGTCTTAAAGCTGCGCCACGCTGGCGGC 3118

DB 418 TCCAGCGACCAAGACCGGCTTGAGCGGAGCGTCTTAAAGCTGCGCCAGCTGGCGGC 477
QY 3119 GGATCAGAAATGATGAGAGCTGCGCTCCAGCCAGAGAGCTGCTAGTGGCCCTGTATTA 3178
DB 478 GGATCAGAAATGATGAGAGCTGCGCTCCAGCCAGAGAGCTGCTAGTGGCCCTGTATTA 537
QY 3179 CTTGAATACCCCGCAATGACCCCTTTTACTGCGCGACCTTCCAAAGGTATATCAGACTC 3238
DB 538 CTTGAATACCCCGCAATGACCCCTTTTACTGCGCGACCTTCCAAAGGTATATCAGACTC 597
QY 3239 TGGCCGATCTGATCATCTTGCACATAGAGCGGCGCAAGCCAAAGTTGCAATTCGGGTGC 3298
DB 598 TGGCCGATCTGATCATCTTGCACATAGAGCGGCGCAAGCCAAAGTTGCAATTCGGGTGC 657
QY 3299 CAATTCGCTAGTAGCTCTCATTTGAGCAGTAGAGTCCCAAGCTTTGCAAGTCCCTC 3358
DB 658 CAATTCGCTAGTAGCTCTCATTTGAGCAGTAGAGTCCCAAGCTTTGCAAGTCCCTC 717
QY 3359 TGTGGTTCATTTGCTGCTTCCAGAGCCACCAACCACTTAGCATCA 3408
DB 718 TGTGGTTCATTTGCTGCTTCCAGAGCCACCAACCACTTAGCATCA 767

RESULT 13
AR510918 767 bp DNA linear PAT 22-SEP-2004
LOCUS Sequence 15878 from patent US 6703491.
DEFINITION AR510918
ACCESSION AR510918
VERSION AR510918.1 GI:52446393
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 767)
AUTHORS Homburger, S.A., Ebens, A.J., Jr., Erickson, C.S., Francis-Lang, H.L.,
Margolis, J.S., Reddy, B.P., Ruddy, D.A. and Buchman, A.R.
TITLE Drosophila sequences
JOURNAL Patent: US 6703491-A 15878 09-MAR-2004;
Exelixis, Inc.; South San Francisco, CA

FEATURES
source location/Qualifiers
1..767
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 14.2%; Score 732.8; DB 6; Length 767;
Best Local Similarity 98.7%; Pred. No. 9.1e-175;
Matches 760; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

QY 2639 CCTGATCAGCTTCACACAGTATCTGGCCGATGAAAGAGCTCACTAGCAGCAACTTAA 2698
DB 1 CCTGATCAGCTTCACACAGTATCTGGCCGATGAAAGAGCTCACTAGCAGCAACTTAA 60
QY 2699 ATGCGTTCTGGACATGTTCCGCAAGATGTTCAATGATACACACACCAAGTGTACTACT 2758
DB 61 ATGCGTTCTGGACATGTTCCGCAAGATGTTCAATGATACACACACCAAGTGTACTACT 120
QY 2759 GTTTTGGACACGGTACCGAAGTATCTGGTTCATGCGAAGAAATTCACGAATGG 2818
DB 121 GTTTTGGACACGGTACCGAAGTATCTGGTTCATGCGAAGAAATTCACGAATGG 178
QY 2819 CTCTTCATCTTGTGACGCGGTGTTTCACAAATTTGGGACCGATCTACTCAATTGGAT 2878
DB 179 CTCTTCATCTTGTGACGCGGTGTTTCACAAATTTGGGACCGATCTACTCAATTGGAT 237
QY 2879 GCACAGCAAGTTTGAAGACGCTACAGGTGCTCAGCAATATTTCCCGACGCGATTGCA 2938
DB 238 GCACAGCAAGTTTGAAGACGCTACAGGTGCTCAGCAATATTTCCCGACGCGATTGCA 297
QY 2939 GCTTAAAGAACTATTGAATCATATGGAATTTCTACTCAAAAGCCCACTACAGAGCGG 2998
DB 298 GCTTAAAGAACTATTGAATCATATGGAATTTCTACTCAAAAGCCCACTACAGAGCGG 357

QY	2999	CATGCCCAATTCGGGCTTCCTTAGGGATCTGGCCAAATCATATCTGTAATAAGAGGACTT	3058
Db	358	CATGCCCAATTCGGGCTTCCTTAGGGATCTGGCCAAATCATATCTGTAATAAGAGGACTT	417
QY	3059	TCCGACGCACGAGCCAGGCGCTGCGAGCGAGCGGTCTCTTAAGCTGGGCCAGCTGGCGGC	3118
Db	418	TCCGACGCACGAGCCAGGCGCTGCGAGCGAGCGGTCTCTTAAGCTGGGCCAGCTGGCGGC	477
QY	3119	GGATTCGAAAGTCGATGGAGCTGCGCTCCGAGGCCAGAGCTGCTTAGTGGCCCTGTATTA	3178
Db	478	GGATTCGAAAGTCGATGGAGCTGCGCTCCGAGGCCAGAGCTGCTTAGTGGCCCTGTATTA	537
QY	3179	CCTGAATACCCCGCAATGACCCCTTTACTGGCGACCTGGCCAAAGGATATACGAGATC	3238
Db	538	CCTGAATACCCCGCAATGACCCCTTTACTGGCGACCTGGCCAAAGGATATACGAGATC	597
QY	3239	TGCCCGATCTCTGCATTCATTCGCACATGAGGCGGCAAGCCAAAGTTTCAATTGGGGTGC	3298
Db	598	TGCCCGATCTCTGCATTCATTCGCACATGAGGCGGCAAGCCAAAGTTTCAATTGGGGTGC	657
QY	3299	CAATTGGCGTAGTAGCTTCCATTGAGCAGTAGCAGTCCCAAGCCTTTGCCAAAGTCCCTC	3358
Db	658	CAATTGGCGTAGTAGCTTCCATTGAGCAGTAGCAGTCCCAAGCCTTTGCCAAAGTCCCTC	717
QY	3359	TGTGGGTCCATTGGCTCGCTTCAGAGGACACACACCAACTTATGACATCA	3408
Db	718	TGTGGGTCCATTGGCTCGCTTCAGAGGACACACACCAACTTATGACATCA	767

	RESULT 14							
	AR501058							
LOCUS	AR501058	1420 bp	DNA	linear	PAT 22-SEP-2004			
DEFINITION	Sequence 6018 from patent US 6703491.							
ACCESSION	AR501058							
VERSION	AR501058.1	GI:52436533						
KEYWORDS								
SOURCE	Unknown.							
ORGANISM	Unknown.							
REFERENCE	Unclabified.							
AUTHORS	1 (bases 1 to 1420)							
TITLE	Homburger,S.A., Ebens,A.J. Jr., Erickson,C.S., Francis-Lang,H.L., Margolis,J.S., Reddy,B.P., Ruddy,D.A. and Buchanan,A.R.							
JOURNAL	Drosophila sequences							
FEATURES	Patent: US 6703491-A 6018 09-MAR-2004;							
source	Exellixis, Inc.; South San Francisco, CA							
	Location/Qualifiers							
	1..1420							
ORIGIN	/organism="unknown"							
	/mol_type="genomic DNA"							
Query Match	12.8%; Score 661; DB 6; Length 1420;							
Best Local Similarity	99.0%; Pred. No. 1.9e-156;							
Matches 676; Conservative 0; Mismatches 5; Indels 2; Gaps 1;								
OY	4465	TTGTACATCGGAGCAGCCGGAACTGCATCACTGGGGGAGGAAGTCTTACAAGAATCTCT	4524					
Db	1	TTGTACATCGAAGCACCGGAACGTGATCATGATGGGGGAGGAAGCTTACCAGAACCTCT	60					
OY	4525	CCGCCGCATGTCGTCAATGATTGGGGAGAGCCCTTAATAGGATTTCCTGC-	4582					
Db	61	CCGCCGCATGTCGTCAATGATTGGGGAGAGCCCTTAATAGGATTTCCTGCATCAGACACACC	120					
OY	4583	ACAAACAGACACAGACGCGGTTGCTTCCTCGGCTTGAGAGGAAGTAGAGAGCGCGG	4642					
Db	121	ACAAACAGACACAGACGCGGTTGCTTCCTCGGCTTGAGAGGAAGTAGAGAGCGCGG	180					
OY	4643	ACATTAAATATATATATATATACACTTAATCTTAATATTTATATACATATATACCATCATC	4702					
Db	181	ACATTAAATATATATATATATATACACTTAATCTTAATATTTATATACATATATACCATCATC	240					
OY	4703	CCTTACTTGTGTAACTGCTCGATGATATATTACAGCAGATGCAAAATGCGCCCGAAGAA	4762					

Db	241	CGTTACTGTGTAACTGCTGATGTCATATATTTGACGAGATGCAATATGGGCCGAAAAGAA	300
QY	4763	AGTCAAGGGCCATGCCCATCTAAATGTGAGCAGAAAACTATTTATATATAACGGGAAC	4822
Db	301	AGTCAAGGGCCATGCCCATCTAAATGTGAGCAGAAAACTATTTATATATAACGGGAAC	360
QY	4823	TAAAGCGAGTAAATCCGCAAAAGTGTACTTAAATTTGTAGCCAACTCCGAAATCCACACTC	4883
Db	361	TAAAGCGAGTAAATCCGCAAAAGTGTACTTAAATTTGTAGCCAACTCCGAAATCCACACTC	420
QY	4883	TCACATCAGTTGTACGTCTTTTACGACCTCCCTACTATTAAACGTCATTTCTGTTAAT	4942
Db	421	TCAATCATGTTGTACGTCTTTTACGACCTCCCTACTATTAAACGTCATTTCTGTTAAT	480
QY	4943	TGTAAGCCTATATACCGCTCTTTATGTGAAACTCAGACGTGTGCTTTTCTACCTTTGTTTC	5002
Db	481	TGTAAGCCTATATACCGCTCTTTATGTGAAACTCAGACGTGTGTCTTTCTACCTTTGTTTC	540
QY	5003	AGCGCACTTGTGATCCTTATTTGCTTCCGCTCAACTAAATCTAATTACTATTATAT	5062
Db	541	AGCGCACTTGTGATCCTTATTTGCTTCCGCTCAACTAAATCTAATTACTATTATAT	600
QY	5063	ACGATTTGCTCTTACAACTGAACATAATTTTGTAAATTAATTGAATACAAAGCCACAC	5122
Db	601	ACGATTTGCTCTTACAACTGAACATAATTTTGTAAATTAATTGAATACAAAGCCACAC	660
QY	5123	AAAGATTTAAATTCATATAAAAAA	5145
Db	661	AAAGATTTAAATTCATATAAAAAA	683

[illegible]


```
QY 4703 CGTTACTTGTGTAACTGCTGATGCAATATATTCAGCAGATGCAATGCGGCCGAAAGAA 4762
    |||||
Db 241 CGTTACTTGTGTAACTGCTGATGCAATATATTCAGCAGATGCAATGCGGCCGAAAGAA 300
    |||||
QY 4763 AGTCAAGGGGCGCATGCGCCATCTATATGTGAGCAGAAAACTATTATATCAATAACGGGAAC 4822
    |||||
Db 301 AGTCAAGGGGCGCATGCGCCATCTATATGTGAGCAGAAAACTATTATATCAATAACGGGAAC 360
    |||||
QY 4823 TAAAGCGAGTAAATCCGCAAGTGTACTAAATTTGTAGCCAACTCCGCAATCTCACTTC 4882
    |||||
Db 361 TAAAGCGAGTAAATCCGCAAGTGTACTAAATTTGTAGCCAACTCCGCAATCTCACTTC 420
    |||||
QY 4883 TCACATCAGTTGTAGCGTCTTTTACAGCTCTTAATTAAGCGTGAATTCGTTAAT 4942
    |||||
Db 421 TCACATCAGTTGTAGCGTCTTTTACAGCTCTTAATTAAGCGTGAATTCGTTAAT 480
    |||||
QY 4943 TGTAAAGCTATATACCGGCTCTTATATGGAACCTCAGAGTGCGTTTCTACCTTGTTC 5002
    |||||
Db 481 TGTAAAGCTATATACCGGCTCTTATATGGAACCTCAGAGTGCGTTTCTACCTTGTTC 540
    |||||
QY 5003 AGCGCACTTGTATCCCTTATTTCCGCTCACTAAATCTAATTACTATTATTAAT 5062
    |||||
Db 541 AGCGCACTTGTATCCCTTATTTCCGCTCACTAAATCTAATTACTATTATTAAT 600
    |||||
QY 5063 ACGATTCTGCTCTTACAACTGAACATAATTTTGTAAATTAATTAAATACAGGCCACAC 5122
    |||||
Db 601 ACGATTCTGCTCTTACAACTGAACATAATTTTGTAAATTAATTAAATACAGGCCACAC 660
    |||||
QY 5123 AAAGATTAATAATCAATAAAAA 5145
    |||||
Db 661 AAAGATTAATAATCAATAAAAA 683
    |||||
```

Search completed: November 29, 2005, 21:13:39
Job time : 16708 secs

Db 815 TCAATGCTCTGGAGACAGACTTTAACAATAAGCAGATTGTGCCACATATATGCAACT 874
 Qy 592 TTCTGGAGATCCCAAGTATATGTGAGAGAGCGGCATCCAAACGCTAGTGGAAATCT 651
 Db 875 TACTTGGAGATCCAAACAGCCAGGTTGAGATGACGACATTAACGCTTAGTGGAAATTT 934
 Qy 652 ACAAGCATGTAGGGGATGATTCGCGCCAGACTTCCTCGATGAGACGATTTCTGCT 711
 Db 935 ACAGCATGTAGGAGAACTGTGAGGGGACAGATCTCAGTAAATA---AAGATTGGCACAGT 991
 Qy 712 CGAAATTTGCTATGTGTGAGCAAAAAGTTGACAGGTCAAAACAGAGGCTTACTGCTAC 771
 Db 992 CCCGGTGAATGTATTTTTCAAAATTGTGATGATCAAAAATCTGGAAAACATGATAC 1051
 Qy 772 CTTCAGCCCTTAAAAACAGATGGCATGTGAGGTGGCTTTGAGCAGAGCC-GACAAATAT 830
 Db 1052 AATCTGCAATGATTAATAAATTTCAGCATGAAAGATTCTGTGATGTGTAACAGACTTCT 1111
 Qy 831 GGGTTGAGGAGCGACCCACAGAGATGATTAAGCGGCACATCACTCGCGCTTTCGTCA 890
 Db 1112 CTGCTAGTTTACATCATCCAAAGGCTCCACCAGTTCTCGAGAAACGTTGGAAATGGAA 1171
 Qy 891 -----TCAGTGGCGCCCAAAACCAATGTGAAGATGTGACCGGTATG 933
 Db 1172 CCACCCCGCGGCTTGTTCATCCACCTTGGATCCAACTTCAGCTGCAAAAAGAGAG 1231
 Qy 934 CCGGCGCGGTACCATGGAATCTTTGCAATAGCTTGGAGGTGGTCCCGCAATTGAACA 993
 Db 1232 CTGCTGCTGTGTGATGAGAGATTTTATAAGCATTTGATGATGTACTGTATGACGA 1291
 Qy 994 TCTTTCACGCTAAGACATGACGATATCTTCAAGCAAGTACTAGTATCATAGTATTA 1053
 Db 1292 TTATTCAGCGCGAGACTTGTAGGATCCATTAACAAATTAGGAAATATATCTGATG 1351
 Qy 1054 AAAACGAGACTGGAGAAAGTGTGATGCTCTCAAGAAATCAGGGCATTTGCTCATC 1113
 Db 1352 ACAAGCATGTGTGAGAGAGAGATTAATGCTTAAAAAAGATTAGATCTTTACTTTT - 1409
 Qy 1114 TCAGCTATCACACTCAGCGCGAGTTTGTGCTGTACAGCTAAAGAAATGTGCTTAAAGT 1173
 Db 1410 -----GGCTGTGCTGTGATGATGATTAATCTTTCAACATTTGCTTTTGGATGAG 1465
 Qy 1174 TCGTGACATCTCTAAGAGAGAACTACGATCAGGTATCCGAGGCGTGATCAACA 1233
 Db 1466 CCTTTAACTCTCTGCTAAGACCTCGGCTCAGGTATGCGGAGGCTGTATCAAGT 1525
 Qy 1234 TCGCTTACATGTCTTAAGACGCTGAGAAATAAAGTCTTCTGCTGAGCATTTTGG 1293
 Db 1526 TGGGGCATCTGTCTATCAGTTCTGGGAAATTAAGTTGACATGAGCTGAAAGCAATATGC 1585
 Qy 1294 AGCACCTGATTAATTTAATACAGAACAGCGGAAGGTATTCAGCTTCCCAATATG 1353
 Db 1586 CAATATCTTTAATTTAATTTCAAAACAGTGCCTCAAAATATATGCAATCTGTGTGAG 1645
 Qy 1354 CTCTGAAGTATATCAATTAAGTATACATGACCAAAAGCTGCTCAAGATCTTACAGCA 1413
 Db 1466 CTGTAGGTATTAATTTGCGACACACATCTCAGGTATATCTGTCTAATACAGCA 1705
 Qy 1414 CTCTGAATCATCAAAAGTCAAGACATAAGTCCACCTGTGAGAGTGTGTGCTGC 1473
 Db 1706 ---ACTGTAAGCTCTGAGTCTGTGCAAGTTAAGAGCGCTGTTTGAATTTTAAATTTGC 1752
 Qy 1474 TCTTCGAGAGTGGCAGACAGAGCGTTGAAAGAAATGCAACGCTAAGAGGACACT 1533
 Db 1763 TTTTACAGAGATGGACAGACATTCATAGAACAGACATATCAGTATTTACTGAAACAA 1822
 Qy 1534 TAAAAAAATCATTTGGCGATGCAACTGCGATGCAAGCGCCCATTTCCAGATACCGCTAT 1593
 Db 1823 TAAAGAAAGGAATCATGATCTGATCTCGAAGCAAGATAGAACCCGAAATATGTTACT 1882
 Qy 1594 GGGCTTTCAGGCGTCACTTTCAGAGCTGGGCGGATCAATATATGAACTTATAGACATAG 1653
 Db 1883 GGGGTTTTCACATGCTTTCAGCAGAGAAAGACCTGTATACCAACCTTTGAGATCT 1942

Qy 1654 CTGCCAGCGCGC 1666
 Db 1943 CTTACCGAAGAAC 1955
 RESULT 2
 US-10-104-047-517
 ; Sequence 517, Application US/10104047
 ; Publication No. US20030236392A1
 GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; FILE OF INVENTION: No. US20030236392A1 full length cDNA
 ; CURRENT APPLICATION NUMBER: US/10/104,047
 ; PRIOR FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER:
 ; NUMBER OF SEQ ID NOS: 4096
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 517
 ; LENGTH: 2524
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-104-047-517
 Query Match 2.8%; Score 141.8; DB 6; Length 2524;
 Best Local Similarity 46.1%; Pred. No. 1.1e-31;
 Matches 768; Conservative 0; Mismatches 852; Indels 45; Gaps 7;
 Qy 77 GAAGCCAGCGACTGTGATGCTTTATCCAGCAATATGCCAAGCGGACATGCTGTGA 136
 Db 357 GAGAGCCCGCAGATGAGTACTTTCGCGCCAGTGCAGCAGAGACGTCGGCGCCG 416
 Qy 137 GGTACAGCTCCCGAGAGATGTGTGACATTCCTTAGG---AGACACAAACTCAATTGT 193
 Db 417 GCTCAGAGTGGCAGAGAGCTCTGCTCTACCTTGGGCCCCGGGCGCATTCGACCT 476
 Qy 194 CTCACGAGCATGAGATTCTCT-----TATTGACGTTTGTATGTCATGCTGACGG 244
 Db 477 GAGAGAGACCTGGCGCGCTAGCAAGACAGTGCAGCGCTCACCGGCTGGGTGGTTC 536
 Qy 245 CAGCCACTTAAATGACACAAAAGTCCCTGAGGCGTTTCCGAGCTAATAAGCAT 304
 Db 537 GACCAACTACCGGGTATCAATTAATGGGATGAAATTTTAAGTCCCTTGTGACAGATT 596
 Qy 305 GGGCAGGATTTAATGACATACAGGCTACCGTTCGCCCATGTGATGATGGCTGG 364
 Db 597 ATCAACAGCTTTAAATCTATGTAGCATGTTATTTAGCTTTAATAGCAGATGG 656
 Qy 365 AGACAGAGGAGACAGTCCGCGAGAGAGCGCAACTTGTGCGGCACTCATGAGCA 424
 Db 657 AGATGCCAAAGACAGAGTTCAGATGAGTCAAGCTGATATTTGAATTAATGATCA 716
 Qy 425 CAGAGTCTTCGCCCCAGGCGCTGATTCGAAAGCTGGCCCATAGCTCTTCAAGCAA 484
 Db 717 AGTAG-----CACCACTATATGACATTTGGGAGCAGTTGGCTTGTGTTAAACCAA 770
 Qy 485 GAAGCGCAAGGTGGGAGAGATTCCTTCAGAGATTTGAAAGCTCTTATGATAGTACG 544
 Db 771 GAATTTTCATCTCGAAGAAAGCGTGTGTCTTTATTTAAACCTTTAACTTTTGG 830
 Qy 545 CACCCAGACCTTATGTTGTGCTGTATATATACACAGATTGTGCACTTCGAGATTC 604
 Db 831 GGCTCAGCCACTGTATCATGCAAAATTAATGATTAACCAATTTGTATCTGTTGAGACTC 890
 Qy 605 CACAGTTAATGTAGAGAGCGGCGCATCCAAAGCTATGAGAAATCTACAGATGAG 664
 Db 891 CAACAGTCAGTGAAGATGTCTCAATATTTGGTATATGAGAGATTTAATGACATGGG 950
 Qy 665 GATTCATTTGGCGCCAGACCTCCGTGCAATGACATGATGTTCCGCGC---TCGAAATTGG 720
 Db 951 AGAAAAAGTAGAGATGATTTTATTAAGAGGAATTTCCCTGCTAGATTAAGAAATGAT 1010

```

QY 721 CTATGTTGGAGCAAAAGTTCGACAGGTCTACAGAGGGCTCTACTGCTACCTTCAGCCC 780
DB 1011 ATTTGGCAATTTGATGAGTGCAGAGTTGACGGCGATATGATTTAGTGTCTGCAAGA 1070
QY 781 TTTAAAAACGATGTCATGAGTGGCTTGGACAGAGCCGACATATTTGGTTAGGG 840
DB 1071 TAAAGCTTCGATGAGTGAAGAAATCACTGATGAAATAGGCCATCATCAGCTGCATCAGC 1130
QY 841 AGGACCCACAGATGATTTAAGCGGCACTACCTCGGCCGTTTG----- 887
DB 1131 CTTCAAGGTTCCCTGCACCTTAAACATCCGAAATCTCCCAACAGTGCAGAGAGCTGG 1190
QY 888 -TCATCACTGCGCCCAAAACCCAAATGTGAAGATGACCGGATGATCGCGCGCGTAAAC 946
DB 1191 TTCAGCAGGTGGCCCTAAGGTTGAGGTGCTCTTAAGAAAGAGGTGCTGAGACAGTTGA 1250
QY 947 CATGGAATCTTTCGAACTTAAGCTTTGAGTGGTCCCGCAATTGACATCTTCCAGCTTA 1006
DB 1251 TGAAGATGATTTTATTAAGCTTTTACAGATGTCCTTCTATTCAGATTTATTTCTAGTCG 1310
QY 1007 GGACATGACGATATCTACAGACAGTACTAGTATCATGATGATTAATAAGCAGAGCTG 1066
DB 1311 AGAATCTGAAGAAACATTAATTAATTCAGGGAATTTTGTCTGATGATTAACATGACTG 1370
QY 1067 GGAGAAACGTGTGATGCTCTCAAGAAAGATCAGGGCATTTGCTCATTTCTGAGCTATCAG 1126
DB 1371 GGAATCGCGTCCCAATGCACCTGAAGAAATTTGATGATCAGTCTTGT-----GCTGAGC 1424
QY 1127 TCAGCCGAGTTTGTCTGTACAGCTTAAAGAAATGTGCTTAAGCTTGTGTGACATCTCT 1186
DB 1425 TGCACAGTATGATGCTTTTTCACATTTTACATTTGATGATGAGGACATTAACCTTTTC 1484
QY 1187 CAAGGAGAACTACAGTACAGGTGATCCGCGAGGGGTGATCAGATCGCCATCAGATGTC 1246
DB 1485 AGCTTAAGATCTTAATCCAGGTGTGAGAAAGCTTGATTAATCTGTAGTACCACTTTTC 1544
QY 1247 TAAGAGCTGAGAAATTAATTAAGTCTTCTGTGAGAAATTTTGAAGACCTGATTA 1306
DB 1545 AACAGTTTGGGAAACAAGTTTGAATGATGATGCGCTGAAGCATTTGATCACTTATTA 1604
QY 1307 TTTAATACGAAGAGCGGAGAGTCAATTGATCCGCTTCCACATATGCTGTGAAGTAT 1366
DB 1605 TCTCGTCCCAATGATGCAAAAGTCAAGCACTTGTGATGTGACGACATCAGATTTAT 1664
QY 1367 CATTAAGTATACATGACCAAGAGCTCAAGATCTACAGACAGACTGTGATCAATC 1426
DB 1665 CATTGCGCACTCATGATACCCAGACTTATACCTTTAATPAACAAGCAATTTGCA---CATC 1721
QY 1427 AAAGTGAAGGACATTAAGTCCACTGTGTGAGTGTGCTGCTCTTTCAGAGAGTG 1486
DB 1722 AAATATAGTTCCCGTGAAGAGAGTTCATTTGAAATTTTATGATTTATTTGTGCAAGAGTG 1781
QY 1487 GAGAGCAAGAGCGCTTGAAGAAAGATCCACCTGATCAAGGACACTTAATAAAATCCAT 1546
DB 1782 GAGAGTCACTTCACTTGAAGAAAGATCAGAGCGCTTGTGTGAATTAATAAAAGGAAT 1841
QY 1547 TGGCGATGAGATGTGATGACAGCGCCCATTTCCAGATTAAGCCTATTTGGCTTTAGGGG 1606
DB 1842 TCATGATGCTGAGCGTGAAGCCAGAGTGAAGGCAAGAAAGCAATATAGGGCTTTAGAAA 1901
QY 1607 TCACCTTCCAGAGCTGCGGATCAAAATATATGGAACATTAAGATAGTCCAGAGCGC 1666
DB 1902 CCACTTTCTGTGGAAGCTGAACATTAATTAATTTCCCTTGAGCCATCTTATACAGAGAG 1961
QY 1667 ATTAGAAGGAAACGAGAGGCGCGGAGAGGAGAACTGTGAC 1711
DB 1962 TCTTCAACTTACTTAACAGTTCGCGAGTGTAGCATCTTCC 2006

```

RESULT 3
US-10-104-047-113
; Sequence 113, Application US/10104047

```

; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 113
; LENGTH: 1908
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-113

Query Match      2.3%  Score 116.6; DB 6; Length 1908;
Best Local Similarity 48.7%  Pred. No. 6,1e-24;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

QY 927 GGTGATGCCGGCGCCGTAACATGGAATCTTTCAGATCTAGCTTGAAGTGTCCGCA 986
DB 353 GGAGGTGCTGAGAGCATGATGAAGATTTTATTAAGCTTTTACAGATGTCCTTCT 412
QY 987 TTGAACATCTTTCACGCTTAAGACATGACGATATCTACAGCAAGTACTAGTATC 1046
DB 413 ATTCAATTTATTTATGTCGAGAACTCGAAGAAATTAATAATCAGGAAATTTTG 472
QY 1047 AGTATTAATAAGCAGACTCGGAGAAAGTGTGATGCTTCTCAAGATTCAGGGCATTG 1106
DB 473 TCAGATGATTAACATGACTGGATGATGCGCAATGACGTAAGAAATTTGATCCTG 532
QY 1107 CTCACTTCAGGATATCACACTCAGCCGAGTTTGTGCTGTACAGCTAAAGAAATTTGTC 1166
DB 533 CTGTGTGCTGAG-----CTGCAAGATATGCTTTTTCACATTTAGCATTTGTTG 586
QY 1167 TTAAGCTTGTGAGACATCTCTCAAGAGAACTACATACAGATGATCCGCGAGCTGC 1226
DB 587 GATGAGACATTAATTAATTTTACAGCTAAGATCTTAATCCAGGTGTGAGAAAGCTTGT 646
QY 1227 ATCAACATCGCCTATATGCTTAAGAGCGTGAAGAAATTAATGATGCTTGTGTCGAGC 1286
DB 647 ATTAAGTACCCACCTTTTCAACAGTTTGGAAACAAATTGATCATAGCCCTGAAGCC 706
QY 1287 ATTTGAGACACTGATTAATTTATACAGAAACAGCGGAGAGTCAATTCGCTTCC 1346
DB 707 ATTTGATCTACACTTTTATATCTGTCGCCAATATGCAAAAGTCAATGCACTTCTGA 766
QY 1347 ACAATAGCTCTGAAGTATATCATTAAGTATACATGACCAAGAGCTGCTCAAGATCTAC 1406
DB 767 TGTGAGCAATACAGATTTATATCTGCGCATACTCATATGCCAGACTTATACCTTTAATA 826
QY 1407 ACAGACACTTGAATCAATCAAGTCAAGAGCAATTAAGTCCACACTGTGTAGAGTG 1466
DB 827 ACMAACAATGCA---CATCAAAATCAAGTCCCGAGAGAGCTTCATTTAATTTTA 883
QY 1467 GTGTGCTCTTTCGAGAGTGGAGAGAGAGGCTTGAAGAAAGATGCCACCTACTAAG 1526
DB 884 GATTTATTTGTTCAAGAGTGGAGAGCTCATTTATGGAAGACATGACGCTTGTGTT 943
QY 1527 GACACCTTAATAAAATTCATGCGATGAGACCTGATGACGCGCCATTCAGATAC 1586
DB 944 GAAACATTAATAAAAGGAATTCATGATGCTGACGCTGAGGCGAGAGTGAAGCAAGAAAG 1003
QY 1587 GCTATTTGGGCTTTCAAGCGTATCTTTCAAGAGCTGCGGATCAATTAATGAACATTA 1646
DB 1004 ACATATCATGGGCTTAAGAAACCATTTCTGTGTAAGCTGAAGCAATTAATATTCCTT 1063
QY 1647 GACATAGCTGCCAGCGCATTAAGAAAGGAAACGAGAGGCGGCGAGAGAGAGAACT 1706
DB 1064 GAGCATCTTATCAAGAGTCTTCAAACTTAATTAAGAGTTCGGAGGTGTAGATCT 1123

```

QY 1707 GGTAC 1711
DB 1124 CTTCC 1128

RESULT 4

US-10-205-219-114
Sequence 114, Application US/10205219
Publication No. US2003013803A1
GENERAL INFORMATION:
APPLICANT: Warner-Lambert Company
APPLICANT: Lee, Kevin
APPLICANT: Dixon, Altheair
APPLICANT: Brooksbank, Robert
APPLICANT: Plimock, Robert
TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
FILE REFERENCE: WI-A-018200
CURRENT APPLICATION NUMBER: US/10/205,219
CURRENT FILING DATE: 2002-07-24
PRIOR APPLICATION NUMBER: GB 0118354.0
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 197
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 114
LENGTH: 5614
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: KIAA0627
US-10-205-219-114

Query Match 2.3%; Score 116.6; DB 6; Length 5614;
Best Local Similarity 48.7%; Pred. No. 1.3e-23;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

QY 927 GGTGATGCGGCGCGCGGACCATGGAATCTTTCGAATCTGAGCTTGGAGTGTCCCGCAA 986
DB 269 GGAGGTGCTGGAGCATGTAAGATGATTTTAAAGCTTTTACAGATGTCCTTCT 328
QY 987 TTGAACATCTTCCACGCTAAGACATGACATGATCTTACAGCAAGTACTAGTATCATC 1046
DB 329 ATTCAAGATTATTTCTAGTCCAGAACTCGAAGAAATTAATAAATACAGGAAATTTTG 388
QY 1047 AGTATTAACACGACAGCTGGGAGAAAGTGTGATCTCTTCAAGAAAGATCAGGCAATG 1106
DB 389 TCAGATGATTAACATGATGCTGGGATCAGCGCAATGCACTGAAGAAATTCATGACTG 448
QY 1107 CTGATTTCTGCTATCATCACTCAAGCCGCACTTGTGCTGTACAGCTAAAGAAATTTG 1166
DB 449 CTGTGCTGCTGAG-----CTGCACAGTATGATTTGCTTTTTCACATTTTACGATTTG 502
QY 1167 TTAAAGCTTGTGACATCTTCAAGAGAACTACAGATCAGAGTATCCGCGAGCGTGC 1226
DB 503 GATGAGACATTTAACTTTAGCTTAAGATCTTAGATCCAGGTGTTAGAGAACTTGT 562
QY 1227 ATCAGCATGCTTACATGCTTAAGACGCTGAGAAATTAAGTAGATGCTTCTGCTGAGC 1286
DB 563 ATTACTGTAGCCCACTTTTCAACAGTTTGGGAAACAGTTTGTATGATGCGCTGAAGCC 622
QY 1287 ATTTGGAGACCTGATTAATTTTATACAGAAACGCGGAAAGTCAATGATCCGCTTCC 1346
DB 623 ATTGTACTACATTTTAACTCTGCTCCCAATAGTCAAAAGTCATGAGCAACTTCTGGA 682
QY 1347 ACATAGCTTGAAGTATCATTAATATATACATGATCAGCAAGTGTCTCAAGATCTAC 1406
DB 683 TGTGACGATCAGATTTATCATTTGCGATCATCATATCCACACTTAATCTTTATA 742
QY 1407 ACAGACACTGTGATCAATCAATCAAGTGAAGACATTAAGTCCACATGCTGTGAGCTGATG 1466
DB 743 ACAGACATTTGCA---CATCAAAATCAGTTCCCGTGAAGACATGCTCATTTGAATTTT 799
QY 1467 GTGCTGCTTTCGAGAGTGGCAGACAGAAAGCGTGTGAAGAAAGATGCAACCTACTAAG 1526

DB 800 GATTATTTTTCGACAGAGTGCAGACTCATTTGGAAGACATGACGCCCTTGTGTT 859
QY 1527 GACACCTTAAAAAAATTCATTGGCGATGCGAGTGCAGCGCCCATTTCCAGATTC 1586
DB 860 GAACTATTAATAAGGAATTCATGATGCTGACGCTGAGGCGAGAGTGAAGCAAGAAAG 919
QY 1587 GCCTATTTGGCTTTCAAGCGGTCACTTTCCAGAGCTGCGGATCAATATATGGAACATTA 1646
DB 920 ACATACATGCTGCTTGAAGAACCACTTCTGCTGTAAGCTGAACATTAATATTCCTT 979
QY 1647 GACATGCTGCCAGCGGCATTTAGAAAGGAACGAGAGCGCGGAGAGAGAACT 1706
DB 980 GAGCCATCTTATCAGAGAGCTTCAAACTTCTTAAGAGTTCTGCACTGATGATCT 1039
QY 1707 GGTAC 1711
DB 1040 CTTCC 1044

RESULT 5

US-10-037-270-324
Sequence 324, Application US/10037270
Publication No. US20030104529A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunding
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: Tillinghast, John
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/10/037,270
CURRENT FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: PC FL_genes Version 1.0
SEQ ID NO 324
LENGTH: 6445
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (164)..(4051)
US-10-037-270-324

Query Match 2.3%; Score 116.6; DB 5; Length 6445;
Best Local Similarity 48.7%; Pred. No. 1.5e-23;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

QY 927 GGTGATGCGGCGCGGTACATGATCTTTCGAATCTTGAAGTGTGAGGTCGCCGAA 986
DB 362 GGAGGTGCTGAGCACTTATGAAATGATTTTAATAAGCTTTTACAGATGTCCTTCT 421
QY 987 TTGAACATCTTCCAGCTTAAGACATGACATATCTTACAGCAAGTACTAGTATC 1046
DB 422 ATTCAAGATTATTTCTAGTGAAGACATGCAAGAAACATTAATAAATCAGGAAATTTTG 481


```

GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
FILE REFERENCE: 784CIP2BDV3
CURRENT APPLICATION NUMBER: US/10/122,851
CURRENT FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: 09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 324
LENGTH: 6445
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (164)..(4051)
US-10-122-851-324

```

```

Query Match      2.3%; Score 116.6; DB 9; Length 6445;
Best Local Similarity 48.7%; Pred. No. 1.5e-23;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

```

```

927 GGTGATCCGCGCGCGTAAACCATGGAATCTTCAATCTAGCTTTGAGTGTCCGCA 986
362 GGAGTCTCTGGAGAGTGTGAAGATGATTTTAAAGCTTTTACAGATGTCCTTCT 421
987 TTGAACATCTTCCACGCTAAGAGACATGACGATCTCAAGCAAGTACTAGATCATC 1046
422 ATTCAGATTATTTCTAGTCGAGAACTCGAAGAACTTAATAATCAAGGAATTTTG 481
1047 AGTGATTAAGAAAGCAGATGGGAAACGCTGTGATCTCTCAAGAAATCAGGCGATTG 1106
482 TCAGATATTAACATGATGGATCGATCGATCGATCAATCAAGAAATTCATCATCTG 541
1107 CTCATCTCAGCTATCACTACAGCCGAGTTTGTCTGCTGACAGTAAAGAAATGTG 1166
542 CTGTTCTCTGAG-----CTGCACAGTATGATGCTTTTTCACATTAACATTTTG 595
1167 TTAAGCTTGTGACATCTTCAGAGAGAACTACGATCAAGTATCCGAGGCGTGC 1226
596 GATGACACACTTAACATTTTCAGTAAAGATCTTAGATCCAGGTGTAGAGAACTGT 655
1227 ATCAACCTCCCTCAATCTCAAGACGCTGAGAAATAACATAGTGCCTTCTGTGAGC 1286
656 ATTAAGTACGACCTTCAAGCTTTTGGAAACAGTTTGTATCAATGGGCTGAAGCC 715
1287 ATTGAGACACCTGATTAATTTATACAGAACAGCCGGAAGTGCATTTGATCCGCTTC 1346
716 ATTGATCTACATTTTATCTGCTCCCAATAGTGCAGAAAGTCAATGCACTTTCGA 775
1347 ACAATAGCTCTGAAGTATATCAATTAAGTATACATGACCAAGCAAGTGTCTCAAGATTC 1406
776 TGTGACACATCAATATTAATTCATTTGCGATACATGATACCAAGCTTAATACCTTATA 835
1407 ACAGACCTCTGAATCAATCAAAATGAGAGACATTAAGTCCACACTGTGTGAGCTATG 1466
836 ACAAGCAATTCGA---CATCAAAATCAGTTCCTGTGAGAGACGTTCAATTTGAATTTTA 892
1467 GTGTGCTCTTTCGAGAGTGCAGACAGAGCGCTTGGAAAGAAATGCACTGATCAAG 1526
893 GATTATATGTTGCAAGATGAGACATCATTTGAAAGAAACATGACGCGCTTGTGT 952
1527 GACACCTTAATAAAATCATTTGGAGATGACAGATGCGATGACGCGCCATTTCCAGATAC 1586

```

```

DB 953 GAAACTATTAATAAGGAATTCATGATCTGACGCTGAGGCGCAGAGTGAAGGCAAG 1012
1587 GCCTATTGGGCTTTTCAGCGGCTACCTTCCAGAGCTGGCGGATCAATATATGGAACATTA 1646
1013 ACATATATGGGTCTTGAACAACACTTCTGATGAGAGCTGAAGCAATATATATTCCTT 1072
1647 GACATAGCTGCCAGCGCCGATTAAGAAAGGAACGAGAGGCGCGGAGAGAGGAACCT 1706
1073 GAGCCATCTTATCAGAAAGATCTTCAAACTTACTTAAGAGTTTGGCAGTGTAGATCT 1132
QY 1707 GGTAC 1711
DB 1133 CTTC 1137

```

```

RESULT 8
US-10-037-270-323
Sequence 323, Application US/10037270
Publication No. US20030104529A1

```

```

GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dunhui
APPLICANT: Wang, Zhiwei
APPLICANT: Tillinghast, John
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: No. US20030104529A1e1 Nucleic Acids and
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/10/037,270
CURRENT FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 323
LENGTH: 6487
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (164)..(4093)
US-10-037-270-323

```

```

Query Match      2.3%; Score 116.6; DB 5; Length 6487;
Best Local Similarity 48.7%; Pred. No. 1.5e-23;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

```

```

927 GGTGATCCGCGCGCGTAAACCATGGAATCTTCAATCTAGCTTTGAGTGTCCGCA 986
362 GGAGTCTCTGGAGAGTGTGAAGATGATTTTAAAGCTTTTACAGATGTCCTTCT 421
987 TTGAACATCTTCCACGCTAAGAGACATGACGATATCTCAAGCAAGTACTAGATCATC 1046
422 ATTCAGATTATTTCTAGTGAAGAACTCGAAGAACTTAATAATCAAGGAATTTTG 481
1047 AGTGATTAAGAAAGCAGATGGGAAACGCTGTGATGCTCTCAAGAAATCAGGCGATTG 1106
482 TCAGATATTAACATGATGGATGAGCTGTGCAATGCACTGAAGAAATTCGATCATG 541

```

```

QY 1107 CTCATTCTGAGTATGACACTGACCGGAGTTTGTGCTGTACAGCTAAGAAATTGTG 1166
Db 542 CTTGTTCTGTGAG-----CTGCACAGATGATGCTTTTTCAACTTAACGATTTGTG 595
QY 1167 TTAAAGCTTGTGACATCTTCAGAGAACTACGATGATGATCCGCGAGGCTGC 1226
Db 596 GATGAGACACTTAAACCTTTCAGTACGATCTTGAATCCAGGCGTTAAGAAAGCTTGT 655
QY 1227 ATCAACCTGCGCTCATGTCTTAAGAGCTGAGAAATAAAGTATGCTTCTGTGAGC 1286
Db 656 ATTACTGTAGCCCTTCCTTCAAGCTTTTGGAAACAAAGTTGATCATGCGCTGAAACC 715
QY 1287 ATTTTGGAGACCTGATTAATTAAATACAGAACCGCGAAGTCAATGATCGCTTCC 1346
Db 716 ATTGATCTTACATTTTAAATCTGCTCCCAATAGTCAAAAGTCATGGCAACTTGGGA 775
QY 1347 ACAATAGCTCTGAAGTATATCATTAATATATACATGACATGACCAAGCTCTCAAGATCTAC 1406
Db 776 TGTGACAGCAATCAAGATTTATCATTCGGCATACCTCATGTACCCAGACTTATACCTTAAATA 835
QY 1407 ACAGACACTCTGAATCAATCAAAAGTCAAGAGACATAAGTCCACACTGTGTGAGCTGATG 1466
Db 836 ACAAGCAATTGCA---CATCAAAATCAGTTCCCTGTGAGAGAGCTTATTTGAATTTTGA 892
QY 1467 GTGCTGCTCTTTCAGAGAGTGGCAGACGAGCGCTTGGAAAGAAATGACACCTGATCAAGG 1526
Db 893 GATTTATTTGTCAGAGAGTGGCAGACTCATTTCAATGAAAGACATGACGCGCTTGTGTT 952
QY 1527 GACACCTTAAATAATTCATTGCGGATGACAGATGCGATGCGCGCCATTCAGATAC 1586
Db 953 GAAACATATTTAAAGGAAATTCATGATGCTGACCTGAGCCAGAGTGGAGGCAAGAAAG 1012
QY 1587 GCGTATGCGCTTTCAGAGCGCTTCAGAGCTTTCAGAGCTGCGGATCAATATATGAACTTA 1646
Db 1013 ACATACATGCGCTTTCAGAAACCACTTCTGCGAGAGCTGAAACATTAATATTCCTT 1072
QY 1647 GACATAGCTGCCAGCGCGCATTAAGAAAGGAAAGAGAGCGCGAGAGAGAGAACT 1706
Db 1073 GACCCATCTTATCAGAAAGAGTCTTCAAACTTACTTAAGAGTTCTGCGCATGTAGATCT 1132
QY 1707 GGTAC 1711
Db 1133 CTTCC 1137

```

RESULT 9

```

US-10-117-722-323
; Sequence 323, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_Fl_genes Version 1.0
; SEQ ID NO 323
; LENGTH: 6487
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (164)..(4093)
; US-10-117-722-323

```

```

Query Match 2.3%; Score 116.6; DB 6; Length 6487;
Best Local Similarity 48.7%; Pred. No. 1.5e-23;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

```

```

QY 927 GGTGATGCGCGCGCGCTAATCAATGAAATCTTTGAAATCTAGCTTTGAGGTGTCGCCGAA 986
Db 362 GAGAGTGTGAGAGCACTTATGAAATGATTTTATTAAGCTTTTACAGATGTCCTTCT 421
QY 987 TTGAACATCTTCCACGCTAAGGACATGAGCATATCTACAGCAAGTACTAGTATC 1046
Db 422 ATTACATTTATTTATGTCGAGAACTCGAAGAACTTAATTAATTAATCAAGGAAATTTTG 481
QY 1047 AGTGATTAATAAGCAGACTGCGGAAACCTGTGATGCTCTCAAGAAATCAGGCAATG 1106
Db 482 TGAATGATTAACATGACTGGGATCAGCGTCCAAATGCACTGAAGAAATTCGATCACTG 541
QY 1107 CTCATTCTGAGTATGACACTGACCGCGAGTTTGTGCTGTACAGCTTAAAGAAATTTG 1166
Db 542 CTTGTTCTGTGAG-----CTGCACAGATGATGCTTTTTCACATTTACGATTTGTG 595
QY 1167 TTAAAGCTTGTGAGACATCTCAAGAGAACTACATCAAGTATCCGCGAGGCGTGC 1226
Db 596 GATGAGCACTTAACTTACCTTACGCTAAGATGATCTTAATCCAGGTGTTAGAGAACTTGT 655
QY 1227 ATCAACCTGCGCTCATGTCTTAAGAGCTTTCAGAAATTAATGATGCTTCTGTGAGC 1286
Db 656 ATTACTGTAGCCCTTCCTTCAAGCTTTTGGAAACAAAGTTGATCATGCGCTGAAACC 715
QY 1287 ATTTTGGAGACCTGATTAATTAAATACAGAACCGCGAAGTCAATGATGCTTCTGTGAGC 1346
Db 716 ATTGATCTTACATTTTAAATCTGCTCCCAATAGTCAAAAGTCAATGCACTTGTGA 775
QY 1347 ACAATAGCTCTGAAGTATATCATTAATATACATGACACCAAGCTGCTCAAGATCTAC 1406
Db 776 TGTGACAGCAATCAAGATTTATCATTCGGCATACCTGATCCAGACTTATACCTTAAATA 835
QY 1407 ACAGACACTCTGAATCAATCAAAAGTCAAGAGCATTAAGTCCACACTGTGTGAGCTGATG 1466
Db 836 ACAAGCAATTGCA---CATCAAAATCAGTTCCCTGAGAGAGCTTCAATTTGATTTTGA 892
QY 1467 GTGCTGCTCTTTCAGAGAGTGGCAGACGAGCGCTTGGAAAGAAATGACACCGTACTAAG 1526
Db 893 GATTTATTTGTCAGAGAGTGGCAGACTCATTTGAAAGACATGACGCGCTTGTGTT 952
QY 1527 GACACCTTAAATAATTCATTGCGGATGACAGATGCGCGCCATTCAGATAC 1586
Db 953 GAAACATATTTAAAGGAAATTCATGATGCTGACGCTGAGCCAGAGTGAAGCAAGAAAG 1012
QY 1587 GCGTATGCGCTTTCAGAGCGCTTCAGAGCTTTCAGAGCTGCGGATCAATATATGAACTTA 1646
Db 1013 ACATACATGCGCTTTCAGAAACCACTTCTGCGAGAGCTGAAACATTAATATTCCTT 1072
QY 1647 GACATAGCTGCCAGCGCGCATTAAGAAAGGAAAGAGAGCGCGAGAGAGAGAACT 1706
Db 1073 GACCCATCTTATCAGAAAGAGTCTTCAAACTTACTTAAGAGTTCTGCGCATGTAGATCT 1132
QY 1707 GGTAC 1711
Db 1133 CTTCC 1137

```

RESULT 10

```

US-10-122-851-323
; Sequence 323, Application US/10122851
; Publication No. US20050239060A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua

```


APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
FILE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2BDV3
CURRENT APPLICATION NUMBER: US/10/122,851
PRIOR FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: 09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: PL FL_genes Version 1.0
SEQ ID NO 323
LENGTH: 6487
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (164)..(4093)
US-10-122-851-323

Query Match 2.3%; Score 116.6; DB 9; Length 6487;
Best Local Similarity 48.7%; Pred. No. 1.5e-23;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

QY 927 GGTATGCGCGCGCGCTTACCATGATCTTTCAGTCTTGGAGTGTCCCGCA 986
DB 362 GGAGGTGTGAGGAGCTGATGATGATGATTTATTAAGCTTTTACAGATGCTCTTCT 421
QY 987 TTGAACATCTTCCAGCTAAGACATGAGCATATCTCAAGCAAGTCTAGTATC 1046
DB 422 ATTGATTTATTTCTAGTCGAGAACTCGAAGAACTTAATTAATAGGAAATTTG 481
QY 1047 AGTATTAAGACGAGCTGAGAAAGCTGTGATGCTCTCAAGAGATAGGCAATG 1106
DB 482 TCAATGATTAACATGATGAGATGAGCTGCAATGACCTGAAGAAATTCATCATG 541
QY 1107 CTCATTTCTACCTTCACTCAGCGGAGTTTGTGCTGACAGCTTAAGATTTGCG 1166
DB 542 CTGTGTCTGAG-----CTGCACAGTATGATTTCTTTTCAACATTTAGATGTTG 595
QY 1167 TTAAGCTCTGAGCATCTCAAGAGAGAACTACAGTACAGTATCCGAGGCGTGC 1226
DB 596 GATGAGAGACTTAACCTTCAAGTATGATCTTATGATCCAGGTGTTAGAGAGCTTGT 655
QY 1227 ATCAACATGCTTACATGCTPAAGAGCGCTGAAGAAATAAGTATGCTTGTCTGAGC 1286
DB 656 ATTACTGTAGCCCACTTTCAAGTTTGGGAAACAAGTTGATGAGCGCTGAAGCC 715
QY 1287 ATTTTGGAGCATGATTAATTTAATACAGAACAGCGGAGAGTATGATCCGCTTCC 1346
DB 716 ATTGTACCTTACATTTTAAATCTGCTCCCAATAGTGAAGATCAAGCACTTGTGA 775
QY 1347 ACAATAGCTGAGTATATATTAATATACATGACACCAAGCTCTCAAGATCTAC 1406
DB 776 TGTGAGAGATCAAGATTTATCATTTGCGATCTATGATCCAGCTTAATACCTTAATA 835
QY 1407 ACAGACCTGTAATCAATCAAGTGAAGAGCATAGGTCACACATGTGTGAGCTGATG 1466
DB 836 ACAAGCAATTGCA--CATCAAAATCAAGTCCCGTGAAGAGAGCTTCAATTTGAAATTTT 892
QY 1467 GTGCGCTCTGAGAGTGAAGAGAGAGCGTTGGAAGAGATGACCAAGCTTCAAG 1526
DB 893 GATTATTTGTGCAAGAGTGAAGCTTCAATTTGGAAGAGATGAGCGCTTGTGTT 952
QY 1527 GACACCTTAAGAAATCATTTGCGATGAGATGAGTGAAGCGCGCATTTCAAGATAC 1586
DB 953 GAAATCTTAAGAAAGGAAATTCATGATGCTGACGCTGAGGCAAGTGAAGCAAGAAAG 1012

QY 1587 GCTTATGGGCTTTCAGGCGTCACTTTCAGAGCTGGCGGATCAATATATGAACTTA 1646
DB 1013 ACATACATGGGTCTTAAGAAACCACTTTCGTGAGAACTTAACATTAATATTCCTT 1072
QY 1647 GACATAGCTGCGGAGCGCGATTAAGAAAGGAGAGAGGCGGAGAGAGAGAACT 1706
DB 1073 GAGCATCTTATCAAGAAAGTCTTCAACTTACTTAAGAGTCTGCGAGATGATCATCT 1132
QY 1707 GGATC 1711
DB 1133 CTTC 1137

RESULT 11
US-10-357-930-59229/c
Sequence 59229, Application US/10357930
Publication No. US20040259086A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endege, Wilson
APPLICANT: Monahan, John
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
TITLE OF INVENTION: HUMAN PROSTATE CANCER
FILE REFERENCE: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
CURRENT FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785,276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/211,314
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/219,007
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/255,281
PRIOR FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 62232
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 59229
LENGTH: 608
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 33
OTHER INFORMATION: n = A,T,C or G
US-10-357-930-59229

Query Match 1.8%; Score 92; DB 8; Length 608;
Best Local Similarity 51.7%; Pred. No. 1.2e-16;
Matches 237; Conservative 0; Mismatches 215; Indels 6; Gaps 1;

QY 232 CATGCTGAGCGGAGCACTTAAGTTGCAAAAGTCCCTGAGCGCTTCTGAGC 291
DB 521 CTGGGAGAACTTCAAGATTTACAGGTGCTGTGCTGGGAGATGAGCATCTGTCCGCC 462
QY 292 TAAATTAAGCATTTGGGAGCGATTTTAATGATACAGCGCTTACCGTTCTGCCATGTGA 351
DB 461 TGTGAGCCCGCTGAGAGATCGTTCAAGCGGAGATCGGCAAGTGTGCTGCCAAGTCTAA 402
QY 352 TCGATGCGCTGGGAGAGAGAGGAGACAGTCCGCGAGAAAGGCAACTTCTGCTGGCG 411
DB 401 TAGACACTTAGAGATGCTTAAGAGCTTGTGAGGAGAGAGAGCTTAACCTTGTCTTA 342
QY 412 ACCTCATGAGACAGAGTGTCTTCCGCCAGCGGTGATGCAAGAGTGGCACTAGCT 471
DB 341 AGATCATGATCAAGGTGCTTAATCCCAAGTATGAGGACAAATGCT-----TGAG 288

QY 472 GCTTCAGACAGACAGCCAGAGTGGCGAGAGTCTCTTACAGCATTTGTGAACGCTC 531
 DB 287 GCTTCAGACAGACAGCCAGAGTGGCGAGAGTCTCTTACAGCATTTGTGAACGCTC 228
 QY 532 TTCATGAGTACGAGCAGCCAGAGTGGCGAGAGTCTCTTACAGCATTTGTGAACGCTC 591
 DB 227 TCATTCGCTCTGAGCAGCAGCTTTACAGCATTTGTGAACGAGTGTGCGACATATATGCACT 168
 QY 592 TTCATGAGTACGAGCAGCCAGAGTGGCGAGAGTCTCTTACAGCATTTGTGAACGCTC 651
 DB 167 TACTTGAGATCCAAAGCAGGTTCCAGATTTGAGAGCATTAACAGCTTATGGAATTTT 108
 QY 652 ACAAGCATGTAGGAGGATGATTGGCGCCAGACTTCCTG 689
 DB 107 ACAGACATGTAGGAGGAGGATGATTGGCGCCAGACTTCCTG 70

RESULT 12:

US-10-437-963-83732
 ; Sequence 83732, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ. ID NOS: 204966
 ; SEQ. ID NO 83732
 ; LENGTH: 4171
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_83034C.1
 US-10-437-963-83732

Query Match 1.1%; Score 54.4; DB 7; Length 4171;
 Best Local Similarity 49.6%; Pred. No. 0.00022;
 Matches 226; Conservative 0; Mismatches 221; Indels 9; Gaps 3;

QY 240 ACGGGAGCCACTTTAGATTGACAAAGTCCCTGGAGGCGTTCTCGAGCTAATTAAG 239
 DB 13 AAGGAGCGCAATTCCTCGCTGCGCCAGAGGCGGCTGAGCGGCTCTCGCGCGCGG 72
 QY 300 CGATTGGGCGCATTTTATGATACAGGCTACCGTTCTGCGCAATGATGATCGG 359
 DB 73 CTGCGCGGCGCACTTCAAGATTCACCTCAAGCCCTCTGCGCGCGCGCTGAGCGG 132
 QY 360 CTGGAGAGCAGCAGGAGCAGTCCGCGAGAGGCGGCACTTCTGCTGCGCGAGCTCATG 419
 DB 133 CTGGGCGAGCGGAGAGCGGCTGCGAGAGCGCGCGAGAGCTGCTTGCACCTCATG 192
 QY 420 GAGCAGAGATGCTTCGCGCCAGAGCGCTGATGCAAGGCTGCGCACTGACTCTTCAG 479
 DB 193 GAGGT-----TTCATCAACGACATCATTTGTGAAGAGCTGGAAGTTATGCA-TGAC 246
 QY 480 CAAAGAGCCGCGAGGCGCGAGAGTCTCTTCAAGCATTTGGAAGCTCTTCATGAG 539
 DB 247 CAAAGAGCTGAGGCGGAGAGATTTGTACGCACTTTCAAGAGCTGTTGCGCTC 306
 QY 540 TAGGAGCAGCAGGCTTATGTTT---CGCGTCTATATACACAGTTTGTGACCTTTC 596
 DB 307 TTGGCTTTTACAGAGCTCCTCTTACAGAGGTTGTTTTCACCTGCTGCAATTTGCTG 366
 QY 597 GGAGATCCACAGATTATGTGAGGAGGCGGCGCATCCAAAGCTAGTGAATCTACAG 656

DB 367 AACGATTGTATCAAGTGTGTAGAGCGCTGCAATCTCATGTATGTAGAGATGTATGA 426
 QY 657 CATGTAGGGGATGATGGCGCCAGAGCTTCGCTGCG 692
 DB 427 AACATGGATCTCAATTTCATGAGAGTTCGACGCC 462

RESULT 13

US-10-437-963-89872
 ; Sequence 89872, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ. ID NOS: 204966
 ; SEQ. ID NO 89872
 ; LENGTH: 4420
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_88591C.1
 US-10-437-963-89872

Query Match 1.0%; Score 53.2; DB 7; Length 4420;
 Best Local Similarity 46.6%; Pred. No. 0.00053;
 Matches 246; Conservative 0; Mismatches 273; Indels 9; Gaps 2;

QY 200 GGACATGGGATTCCTTATGAGCGTTGATGCGATGCGGAGCGGAGCCACTTTAGAT 259
 DB 117 GGAGTACGCGCGCTGTGTGACACTGATGATGATGATGATGATGATGATGATGATGAT 176
 QY 260 TGCAAAAAGTCCCTGAGAGCGCTTCTCGAGCTAATTAAGCGATTTGGAGCGATTTTAA 319
 DB 177 CGGCGAGGCGCGCTGACAGCGCTCTCCCGCGCGCGCTGCGCGGAGGAGCACTTTAA 236
 QY 320 TGCATACAGCGCTACCGTTCTGCGCATGTGATGATGATGATGATGATGATGATGATGAT 379
 DB 237 GATTCACCTGAAGCGCGCTGCTCCGCGCGCGCGCTGCGCGCGCTGCGCGGAGGAGGAG 296
 QY 380 AGTCCGCGAGAGGCGGCACTTCTGCTGCGCATGATGATGATGATGATGATGATGATGATGAT 439
 DB 297 CGTCCGAGAGCG 350
 QY 440 CAGGCGCTGATGACAAAGCTGCGCACTAGCTGCTTCAAGCAAGAGCGGAGGAGTGC 499
 DB 351 GACATCATAGTGAAGAGAGCTGGAAGTATGATGATGATGATGATGATGATGATGATGAT 410
 QY 500 CGAGAGTCTCTTCAAGAGATGTGAAGCTCTT---CATGATGAGGAGCGGAGAGCT 556
 DB 411 GGAAGAGTTTGTAGTACAGTGGCAACCGGCTGGGCTTTTGTCTTCAACGAGCTCC 470
 QY 557 TGTGTGCGCTGTATATATCAAGAGTGTGCACTTCTCGAGATCCAGATTTATGT 616
 DB 471 ATTGACAGAGTTCTTCTTCACTGCTCTTGAATGATGATGATGATGATGATGATGATGAT 530
 QY 617 GAGGAGGCGGAGCTCAAGAGCTAGTGAATTTACAGAGATGATGAGGATGATGATGATGAT 676
 DB 531 TCGAGATGCTGTATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 590
 QY 677 CCCAGACTCGCTGCGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 724


```

QY 530 AAGALEREEREGGG-----GGTGTCTGTATETRTYRSRI 564
Db 329 SYOKSLQYLYKSSGVSASLPOSDRSSSSQESLNRPFSSKMTANPSTVAVRVSAGSSKA 388
QY 565 GRTGTGLOKPPRMSRSIAVDTAQAORAKVAQYTLVRSORRKPFGPNNSNOAS-----617
Db 389 SSLPESLO-----RSRSDIDVNAAGAKAH-HAAGOSVRRGRIGACALNAGSASLEDT 441
QY 618 ---MTGAAASGLPRPLNSNSGGTPATTPGVTPEPRGRAG---VSOSQGRSRTSP--669
Db 442 SDRKIDGTASEDGRVRAKLISA-----PLAGMGNAKADSGRRTKWSVQSQGSRSRSGR 496
QY 670 ---STKRDQYGGYGNVYRGATGAI PKAASGIPRSTASRSTSPTR---SGGGLMKSMY 723
Db 497 VLTJTALSTVSSGQVRLVNSASA--QKRSKIPRSGGSRASPSRLSVASSRIPRPSV 554
QY 724 STGA---GSRTPRRNNVPSAP-----ARLLAQ 750
Db 555 SGGGSRASRSRSDTSFVRSFQPLASRHSRSTGALYAPRVYCASGPGVIGISGSRSS 614
QY 751 SREAEHTLVGDDQDPVYSGDYMRSGMRKLMGRDESDDIDSEASSVCSRSRSPSS 810
Db 615 SVSAMRVLNTGSD--VEEAVALDKKPARRR--YESYGMHSDDDANSDASSACRSRYSVS- 670
QY 811 YTRGNKNYSLSGSHTRLDMSTQAPRPDITITIQFCASTMSRKGGLISLTQYLADGK 870
Db 671 ---RNGSI-----PTYMRQTEDEVAELNRCASSNWSERKGLGLQMLNKNOR 715
QY 871 ELTQOOLKCVLMDRKMFMMDTHK--VYSLFIDVTETELIVANETSNNGSSCLTRLFNK 929
Db 716 TLSVELKRLCEITTRKFPADPHGRVSMFLETLVDFIQYAKDL-QDMLFVLTLLQKK 774
QY 930 LGTDLNSMSHSKIWKTQVNHVEYPTOLQKELFRISDSOTPTTTRTALIRFLTDLA 989
Db 775 MGADLVSSVQAKYQKALDVTRESFPNDQFNILMRFTVDQTPSLKVKVAILIKYETLA 834
QY 990 NTYKSSDFPSD--QSOACERTVUKLAQLAADQSMELRSARCLVALYNLNPQNTLL 1047
Db 835 ---KQMD-PGDFINSEBTRLAIVGRVITWTTEPRKSDVRKAAQSVLISLELNPBEFTML 889
QY 1048 LADLPKYODSARSCISHMRROSQSCNSGANSPPSSPLSSSPK-----PIQSPSVG 1100
Db 890 LGALPKFPQDQATYLLHNHLR---NTGNGTQSSMGSEPLTRPRSRAMNSPILSTPNT 945
QY 1101 PPSALQSHHQLSISTSPRSRQSSVE-----OELLFSSELDIOHNIQKTSEE---IRH 1151
Db 946 SQNTLSPSAFQDYENNNSEDIYSLRGVTEAIQNFSPRQEDMNEPLKRDCKKDDGDM 1005
QY 1152 CFFGQYOTALAPNGFNHLYQHDQGOODSCASLSSNSKTOSSANTTQSN---TPBSATM 1207
Db 1006 CGG---PGMSDPFAGGADATDSQOTALDNKASLLHSMPTHSSPRSRDYNPNYSDSISPF 1061
QY 1208 RLDLLEBERTQKNSPTDAKVTIVSINMAENGELILASNLMESEVVRVALTLTKQOPV 1267
Db 1062 NKSLKAMDEDDADQRPD-----LSIDHSD---LVALLKFLS-----NNHERV 1104
QY 1268 ELLOTSLTNIGICIKGNCELPNKGRFSIMMLNLILAEHTDVVIAGLHVSXKIMSNNK 1327
Db 1105 EERKIALYELMLKQESFSVWDEHFKTILLLETLGDKERTIRALALKVLEIRLHOP 1164
QY 1328 MRNHHMFLELLIKTIQCYOH-SKEALRDISMIRPLASPLDLSINIVNPIATGEF 1386
Db 1165 AR--FKXVAELTYMKTLEAHKDPKHEVVRSAEBAVLAITSISEQCIKYLCPITIQIADY 1222
QY 1387 PTNLCAIKILLEVENHGESETDAHLDIVPNNLARSADDTQSWRKAAVFIYVLGVLG 1446
Db 1223 PINLAAIKMQTKVERSKETLNLILPEIMGLIQGIDNSESRSKACVFLVAHVAVIG 1282
QY 1447 EEEKYKPLSVLNPBKVRLNLYIEKORNCISGGSSSTKNSAAS 1490
Db 1283 DE-LKPHLSQLTGSKMLNLNYI-KRAQTGSGGADPTTVDVSGG 1324

```

```

RESULT 2
US-10-104-047-2297
; Sequence 2297, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OR INVENTION: NO. US20030236392A1el full length cdna
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104, 047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 2297
; LENGTH: 916
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2297

Query Match 16.0%; Score 1214.5; DB 4; Length 916;
Best Local Similarity 32.8%; Pred. No. 1.5e-72;
Matches 321; Conservative 180; Mismatches 322; Indels 157; Gaps 29;

QY 9 LDGFIQMKADMRKVRVQLAEDLVTFELSDPTNSIVC-----TDMGFLIDGL-MFWLGSHE 63
Db 5 MESCIAQVLOKQVKGKRLQVQGLIDYFSDQKSADIEHDQTMLDKLVDSGLATSWVNSNY 64
QY 64 KINOKSLEAFSELKELKSGDFNAYTATVLPVHYIDRLGDSRDYVREKAQILLRDLMEHRVL 123
Db 65 KVVLLGMDLISLALVTRLQDRFKAQIGTVLPSLRIDRGADKDSREDDQILLKIMD-QAA 123
QY 124 PQOALIDKLATSCFKHKNKAVREEFLOTIVNALHEXYTQLSVRYVYIPVYCALLGPTVN 183
Db 124 NPQYVMDRLMG-G-FKHKNPRTREGICLCIATLNSAGAQTLTFLSKIVPHCNLLGDPNSQ 182
QY 184 VREAAIQTVLEYIKYHGDRLRPDLRMDVPSKMLMEQKPEQVQVQEGILLPSALKNTN 243
Db 183 VRDPAALNSVEIYIRHGEVRADLSK-KGLPQSLNVIITKPEYVQSGNMIOQA-----236
QY 244 GNGVGDEADNIGLARPRPMIKRPLHSVSSSLRPPKPNVDVTG-----288
Db 237 -NDKNPDEDSVD-GNRPS-----SASSTSKAPPSRRVAVNGCTTRRLGSSTLGSK 286
QY 289 -----DAGAVTWSEFSEFEVVPQNLIFHADMDDIYQOVLIYISDKADMEKRYDALK 342
Db 287 SSAKSGAGAVDEEDFIKAFDDVPVQIYSSRLDESINKIREIISDDKHDEQVNAALK 346
QY 343 KIRALLILSYHPOPOVAVQLELSLSPVDILKEBELRSQVIRACTTIAYMSTLANKLD 402
Db 347 KIRSLILLAGAEYDNFQ-HRLLDGAF-KLSAKDRLSGVVRACITTLGLSLSVLGNKFD 404
QY 403 AFWCSILHEILNLIQSAVIVASASTIALKYIITKYTHAPKELIKIYVDTLNQSCKDIRST 462
Db 405 HGAEMALPITPHILIPSAKIMATSGVAVRLLIRHRIIRLLIPVITSNCT-SGSVAVRR 463
QY 463 LCEMLVLEFEWQKALEBNATVLRDLTKKSIQDADCDARRSRVAYVMAFRHFPELADQ 522
Db 464 CFEFDLLLOEQWTHSLERHISVLAETIKKGJHDADSEARIEARKCYMGFHSFSEAEH 523
QY 523 IYVTLIDIAQORALEREEREGGGGGGTGTCGTAPEPRTYRSIRGTGCTQKPPSRSTS 582
Db 524 LVHTLESSYQKALQSHLK-----NSDSIVSLPQSDRSSS---SQESLNRPLSARSGPT 574
QY 583 AVDTAAQAKRAQATYLVSRQKPLGPNNSQASVTGAASGSLRPRPLNSNSGGTP--640
Db 575 GSTTSAASIVSTKSVSTTGSLSQR-----SRSDIDVNAAGAA--KSVYSSSGGTTPTS 624
QY 641 ---ATTGVS-----TPRPRGAG---V 657
Db 625 SAALPLPGSVASLESRHMEDMEYIGLDGSRIRTRRQSSGSAATVASTDNGRSPRAKV 684
QY 658 SQSGPSRSTSPSTKLQDYGGI-GNYVYG-ATGAI PKAASGIPRSTASRSTSPRSOG 715

```

Db	685	SGSQGCSRSSSPFGKLLGSGYCGLTGGSSNRPPVTPSESEKRSKLPFGQGCSERTSPNRIG-	743
Qy	716	GLMKRSMYSTGAGS-RRTPERNNPVAPSAFARLLAOSREAHENHLGVGDGQPDVYSDIM	774
Db	744	-----LDRFPGIQGPRIQGSVAM-----RVLSTSTDEAAV-----ADAL	779
Qy	775	RSGGNRMGRKLMGRDESDIDSEASSVCSERSFPDSSYTRGNKSNYSLSGSHRLDWSTOR	834
Db	780	EK-PVRRRREPYGMVDDDANSSASSVCSERSYGS--RNGIGPHY-----L	832
Qy	835	APFDIETIITQCASTHSESRKGLLSIQYLDGKELTQOOLKCVLDMFRKMFNDHTK	894
Db	823	RQTEVAEVLNNCASSNWSERKEGLGLQNLKSGQTLRSVELKRLCEIITRRFADPHSK	882
Qy	895	-YSLFLDVTETELIVHANE	913
Db	883	RVFSMFLETLVDVFIHKKD	902

RESULT 3

```

US-10-104-047
; Sequence: 2487, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2487
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2487

```

Query Match 12.1%; Score 920; DB 4; Length 723;

QY	5	KPSDDGIGIQQMPKXDMRVKQJLAADIVTFU--SDOTNSIVCTDMEFL--IDGLMPWTG	60
Db	2	EPREMEYICAOYQOQXDVGGRLDVGOCELLYLCPAIGALSDLEEDLGRKTYDALGCMVGS	61
QY	61	SHFKIAQSLDAFSELLIKRLGSDFNAYATATVPHVIDRLGSDRTVREKAQILLDLMEH	120
Db	62	SNRYVSLMGLLEILSAFVRLSTREFSKYAMVIVALIDEMGDAKDKRBEAQTLTIKLMD-	120
QY	121	RVLPQQAIDKLATSCFQGNKNAKVEEFLQOTVNAIHEHGTQQLSVRYIIPVCAILLDP	180
Db	121	QVAPRPMTIWEOLA-SGFHKKNFRSNGVCLCIETLNIIFGAQPLVYSKILIPMLCLFSDS	179
QY	181	TVNVREAAIQTLVEIYIKVHGRLRPDLRRMDVPAKSLAMEOKPQYQOEGILLPSALK	240
Db	180	NSQVRDAAILLAVEIRYHVGEEKVRMDLYKR-GIPARLEMIIPAKDEVOSSGGMILSYCK	238
QY	241	NTNGNGVGLDEADNIGLRERPT-----RMIKRPLSAVSSSLRPKNVNDVY--	287
Db	239	DKS-----FDDEESVD-GNRPSSAASAKKVPAPKTSGNPANSARKPGSAGGPKVGASKE	292
QY	288	GDAGAVTMESFESSSEFVVPQLNIFIAKMDDIYKQVLVVISDNADWEKVDALKKIRAL	347
Db	293	GGAGAVDEDDDEIKAFSTDVPSIQIYSSRELEETLNKIRILISPDKDDMOQRANALKKIRSL	352
QY	348	LI-----LSYHTQPOFAVQQLKEIISLSEFVILKEELISQVIREACITITAYMSKTLRKRLAF	404
Db	353	LVAGAAGQDCEFGHRLRLDGALKLS----AKDLASQVREACTIVAAHSTYLGNKFPDHG	407
QY	405	CWSLLEHLINLIQNSAKVIVASASTALKYIILKYTHAPKLKIYDTLQNSKDIRSTLC	464

```

Db      408 AEAIVPLFENLVNPSAKMATSQCAIIFIIIRHTHVPRILIPILTINCT--SKSVPRVRRST 466
Qy      465 ELWVLLFEEMOTKALERNATVLRDTLKISIGDADCDARRHSYYAWAFRRHPPELDQY 524
Db      467 EFLDILLOEMQTHSIERHAAVLVETIKKGIHDADAEARVEAKTYWGLRNHFPEGEAETLY 526
Qy      525 GTULIDIAORALEREKEGGG-----CGTGTGCTAETTR 559
Db      527 NSLEPSSYOKSLQTYLNSSGSVASLPQSDRSSSSSGESLNRPPSSKWTANPSTVARVSA 586
Qy      560 TVSRIGRPTGLOKPTPMSRSISADVTAAGRAKRAYOTLYLSROKPLGPNNNSNOAS-- 617
Db      587 GSKKASLPSGLQ-----RSRSDIDVAAAGAKHHAAGCVRSGR-LGAGALNNGSYA 639
Qy      618 -----MTAAASGSLPRPRLNNSGCTPATTTGVSYTPRRGRAG--VSOSDQGSRS 666
Db      640 SLEDTSDKLDGTASDEGDRVRAKLSA---PLAGWGNKKAUSRGRSRTKMWVSOSQPSRS 694
Qy      667 TSPSTKLRDQYGGYIGNRYRGAATGAIKKKASGIPRSTAS 705
Db      695 GSPGRVL-----TTTIASTVSSGQRVLVNS 720

```

RESULT 4

```

US-10-437-963
: Sequence 192355, Application US/10437963
: Publication No. US20040123343A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Kovacic, David K.
: APPLICANT: Zhou, Yihua
: APPLICANT: Cao, Yongwei
: APPLICANT: Wu, Wei
: APPLICANT: Boukharov, Andrey A.
: APPLICANT: Barbazuk, Brad
: APPLICANT: Li, Ping
: TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53221)B
: CURRENT APPLICATION NUMBER: US/10/437,963
: CURRENT FILING DATE: 2003-05-14
: NUMBER OF SEQ ID NOS: 204966
: SEQ ID NO 192355
: LENGTH: 1366
: TYPE: PRT
: ORGANISM: Oryza sativa
: FEATURE:
: OTHER INFORMATION: Clone ID: PAT_MRT4530_88591C.1.pep
US-10-437-963-192355

```

Query Match	8.8%;	Score 667.5;	DB 4;	Length 1366;
-------------	-------	--------------	-------	--------------

```

Qy      50  LIDGIMWLGSHPKIKAKSLEAFSELIKRLGSPFNAYATVALPHVIDRLDSDRDTVEK  109
Db      44  LVDTCMILIDANFRVAQGGIQLISAAVAAGHFKIHNAIVAAVERLDGKQPVREA  103

Matches 327; Conservative 280; Mismatches 614; Indels 303; Gaps 58.

Qy      110 AQLLLRLDMEHRVLPQALIDKLTATSCFKHNAAYREEPIQTIYNALHEYSTQQLSV-  168
Db      104 ARQLLLITLME--VSSPTLIVERAGSVAMTHKSWKRVREEFVATVAAGLFASTELPFLORV  161

Qy      169 YIPPVCALLDDPTNVNBEAIIQTLVEIYKGVGRDLRLRMDVVPASKLAMELQKFOV  228
Db      162 LLSFVLQWMDSNOSVADAAIYCTIEENYTHGSGFHELDQ-HNLPYMLREINSRLERI  220

Qy      229 KQEGLLPSALKNTNGNGVGLDEADNIGLERPTMIKRLPLHSAYSSSLREKPVNDVT-  287
Db      221 EPK-----VPSDGNIMQYKAAVESRSVSNPKR-----GSPRTKSTPRESTL  262

Qy      288 --GDAGAVNTPESSEFEVVPQNLNIFAKQ-MDDIYQVLIISDKRADWEKRDALKKI  344

```

```

Db 263 FGGDT-----DITEKPEPVRVHSEKELIREBEKIAATLVPEK--DMSVRIANMORI 312
Qy 345 RALLLSYHTPOPOFVAVOLKELSLFVILKEEL---RSOVIREACITIAVMSKTLRNL 401
Db 313 EALVYGADIDPSPFLML-LKOL---VPLSTQSLDRSSIVKQCHILANVLSKELDLDF 367
Qy 402 DAFCSLIEHLINLIQNSAKVIASASTALAKYITKYTHAPKLKLYTTTLNOSKSIDRS 461
Db 368 EPCAEFLFPMLEFKLVITLVIAESADTCIKITLNCIKISRLPRLADTAKNDSAVARA 427
Qy 462 TLCELMVLLFEEW-OTKALERNATVLRPLKKSIGDA-----DGDARHRSYAVAMARR 514
Db 428 RSCETALILEWADAPETORSADLYEDILKCCADAMSDTITNDEGVHVKRYSPPSRE 487
Qy 515 HPEPLADQIYGLDIAORALEREEREGGGGTGTGTAPETRRIVSRIGTPTLOKP 574
Db 488 RF-----VQPSRSLSH-----ASGTSALGYTS----- 510
Qy 575 TFSMRISIAVD-TAAAPAKVRAQYTLYSROKRPFGPNNSQASMTGAAGSLPRPLN 633
Db 511 -----AIYAMDKTAISSDSFSSMTLRLSQKTVG--RSSERSLESVLSNKEKVSAD 563
Qy 634 -----SNSGCTPATTPGSVTPRRRGAGVSQSQPSR-----STSPSTKLPOYGGIGNY 684
Db 564 HLSLONSALLDSSVPSIT-----NASARNGSRLLSEMTTQCTERKSRSPYLGNIS 615
Qy 685 RGATGAI-----PKKASGIPRSTASRETSPTSSGGMLMKRSMVSTGASRRTPERRNVR 740
Db 616 SESMTSLSPFRRLERPOEGGRNDESDIRS-----TRRFPOQNVVD 660
Qy 741 PSAPARLIAQGRAEHTLGVDDGPDYVSGDYMSGMRGKLMGDESDDIDSEASS 800
Db 661 -----MPYRDAIHR-----DSHNNHVP--NFORP--LRLKQVMSR-----AS 693
Qy 801 VCSERSPSSYTR-GNKSNY-----SLG-----SGHTLMDSTOAPRPDIETIIQFAS 849
Db 694 ASINHSFDDSOYQSGVGYTDALASLSDALSEGLSPSSDVMYKSAFEFRNLIQ----- 749
Qy 850 THWSEKDLISLTQYLADEKELTQOOLKCVLDMFKRMFMDHTKYVSLFLDYVETELV 909
Db 750 -----QCGQIGIEIT-----QNFPEKWKLFRRHLDDBHKNVAQAASSTLAELIPA 794
Qy 910 HANETSRNGSSCLTRLFNKLTGDLNMSKIMKTLOVHEHYFTQOLKELFRITIDS 969
Db 795 -CKPFESYVERILPYVFSRL-IDPKELVKKPSSSTLIVGRTYADMLLPAVSL-DE 851
Qy 970 TOTPTTKRIALIRPLDLANTYCSDSPSPSQACERTVUKLAQLAADKSMELRSGA 1029
Db 852 QKSP--KAKLAVLEFANKSFSKTYVDSGYSN-SGFLKMLSKLAPL-VHEGAKLKEAS 907
Qy 1030 RSCVALY-----NLNTPQMTLLADY-----PKYQDASARSCISHMRROQ 1072
Db 908 ISGIISVSHFDSTAVLNFIILSVSEQNLRLAKOYTPRIEVLVLYLQSKORPRPX 967
Qy 1073 S---CNSGANSBPSSPLSSSPKPLQSPSVGFAS---LOSHHQLSISTSPSRSS 1125
Db 968 SYDQADYGTSSBEDGYALASKSPF---GRYSSSSLDABGKMMNSVOESTPRNAPMA 1022
Qy 1126 VEOELLESELDIONIKTSEEIRHCFGGQYQTLAPANGFNGLQYHDQGDQSCASLS 1185
Db 1023 -----RTTSDMSIDH-----TSOSIELDTGSEVLTL-----RSRE 1052
Qy 1186 SNSKTQSSANTQOS--NTPESS-----ATWRLD-----NLERERTQNAK 1222
Db 1053 SKNNSTSLVETARSVPNYPEKTDAPLDEDTAISTRLDLSHRAADGHNAAVQSTAEENVQ 1112
Qy 1223 SPTDADAKYITVYSINAEENGELILASNLMB-SEVURVALTTDQOVELLQTSINTLGICI 1281
Db 1113 EGDIAVVKLSIKITTLHADNELISIPQLHQISNGTEVSLERKEBALQOLVKASVDN----- 1167
Qy 1282 KGCNCELNPKFPRSRIMRLNLLEAHTDUNVYAGLHVLSKIM--RSNKKRHHNMWFLELI 1339
Db 1168 ---DISIAKTFNOLITLVLEVLDSDSSTREIALSLVAEMLNQSGAMBS---ISIV 1220

```

```

Qy 1340 LKTIQCYOH-----SKEALRIDSMTIPRLAPSLPLDLISINYNPVATGSEFTNLCAIK 1394
Db 1221 LEKLIHTKDMVAKISNEANQCLNVLAKYDPF-----RCIAVVPFLVSDDEKTLVVCIN 1276
Qy 1395 ILLEVTNHOSEITDAHLDVFPRLARSADDTQSMKKAIVFCIVKLYFVLGEEKVPKL 1454
Db 1277 SLTLVGRLSSEBELMANOLPFLPLPFLPFAFSNOSPQADAKYVFLVDIYIMLGKAFV-PYL 1335
Qy 1455 SVLNPSKYRLNLYIEKORNCISG 1478
Db 1336 EGNLSTQRLVLTIVYANRISQAFSG 1359

RESULT 5
US-10-437-963-186215
Sequence 186215, Application US/10437963
Publication No. US2004012343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Mu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437, 963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 186215
LENGTH: 1308
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_83034C.1.pep
US-10-437-963-186215

Query Match      8.2% Score 624.5; DB 4; Length 1308;
Best Local Similarity 21.2%; Pred. No. 1.5e-32;
Matches 326; Conservative 279; Mismatches 511; Indels 359; Gaps 60.

Qy 62 HFKAQKSLAEFSELIRKLGSDFNAYATATVLPVHIDRLGDSRDYREKAQILLRDMHR 121
Db 8 NFRVAQGGQLALSAAVLADGDFKILNALVPAVERLGGKQVDRDAARQLVTLWE-- 65
Qy 122 VLPPQALIDKLATSCFKHAKVREEFLOTIVNALHBYGTQQLSV-RYIIPVYCALLGP 180
Db 66 VSPFTIIRAGASYAMTHKSMRVREEFRTVATVGLFASTELPLOBVLLSPVQLNDL 125
Qy 181 TVNREAAIQTIVEIYGVGDRLRPLDRMDVPAKSLAMLEKFPDQVQOEGLLPSALX 240
Db 126 NOSVRDAISCIEBMYMWSQFHEELOR-HNLPSYMLKQINSGLDIE-----PKA-R 177
Qy 241 NTNGNAGVGLDEADNIGLREPTRMIKRPLHSAYSSSLPRKNVNDVTGDAGAVTMESES 300
Db 178 SSDG-----ARMQKYIER-----SRSL-----ADATCTGQDSQIT 208
Qy 301 SFEVVPQNLNIFHAKMDIYKQVLVITS--DKADMEKRDALKKIALLLSHTPQOF 358
Db 209 EKPYEP-IRVSEK--ELIREMEKIASALDPEKDMIRIAMAQRIALVYGALIDYPSF 264
Qy 359 VAVQKELSLSPVILKEEL---RSOVIREACITIAVMSKTLRNLDAFCWSLIEHLINL 415
Db 265 LTL-LKOL---VPLAQLSDRRSSIVKQY-----LFLK 294
Qy 416 IONSAKVIASASTALAKYITKYTHAPKLKLYTTTLNOSKSIDRSITLCELMVLLFEEW- 474
Db 295 VVITLVIAESADNCIKITLIRNCKVSRILPLADTAKNDSAILRACSEYALILEYMA 354

```


QY	475	QTKLERNAVATLPTLTKKSIDGADCDARRHSHVAWFRHFPLEADQIYGTLDIAAORA	534
Db	355	DAPETQARSADLYEDDIKCCVADAMEVAATATASCRTMTKTPMPESSRLPMSFPDAVORI	414
QY	535	LEREREGGGGGTGTGTGTABPTRTVSRIGTPTGLOK--PTPSMR-----	579
Db	415	INDE-DGG-----LOKRPSPSLREKGVOLSHASSHA	445
QY	580	-----SIGAVDRAAQRAKAVRQVLYSR-----QKKPLG--PNNNSQMSMCA	622
Db	446	SGTHLAGYSTAIYAMDKSAA-----ISSESSLSRSHLLSQSKKIGRTAERPIEVLSSSK	501
QY	623	ASGSLPRRLNLSNGCGPATTPGAVTPPRGAGVQSQP-----GSRSTSP--STKLUD	675
Db	502	QKVABIEBLKGNSTLDDSLSPSTINNTNGSRLVDYTNPHVANKERSRSYLLSSSES	561
QY	676	QYGGIGNYRGATGAIP-----KKAAGIPRSTASNETSPTRSOGGLMKRSMYSTGASGR	731
Db	562	ISGSSLPRARSSSGRSPYGSTMBEENDTWSTRBMQMDHNYLDMTRYRDAHSRHLNHQ	621
QY	732	TPERNPPYRPAAPARLLAQSRBAHTLGVGDGQPDVYSGDYMNSGGMKMRKLMGRDES	791
Db	622	VPHFQRPRLKQVARSATSASR--HSF--DDG--HISNDM-----	655
QY	792	DDISEASSVCSERSFDSSYTRGNKS--NYSISG--HTFLDMSTQAPARDIETIIQFAS	849
Db	656	-----SRVYDGPSTISDALSGLSASBDVARVARTANFQITLL-----	693
QY	850	THMERKDGILSTQYLDADGKELTQOOLCKVULDMFKKFMOTHTKYVSLFDTVTELYV	909
Db	694	-HQOK-----GIQVWQNEKMKLFFRLYDDPHHKVAAARSTLADIIPA.739	
QY	910	HANETSRNGSSCLTRLFNKLGTDLNLSMHSKIMKTLOVHEYFTQQLKELFRISDS	969
Db	740	-CKQFESYVRIIIPYVFSRL-IDPKELVROPCSGSTTELEVGRTPIDTLPALVSL-DE	796
QY	970	TQPTTTKRIALLPLDNLANTYCKSSPFPDQOAGCRVYLKLAQLAADQKSMELRQ	1029
Db	797	QRSP--PAKLAIVLEPANKSPFRKYVDSGEGYAN--SGFLYLMJSLKAPL--IHKNAKLETS	852
QY	1030	RSCVVALY--NINTEPOMTLLADLPKVYDQDARSCTHSMRR-----OSQSCNGAN	1079
Db	853	ISGIATVYSHRDTAVNLFIAMLSIEBQNLVRALKQVTRPIEVDLVNYLQSKKERSPK	912
QY	1080	SPSSSPSSSSP-----KPLQSPSYGPASLOSHHQJLSISTSPRSQSSVEQELLFSS	1134
Db	913	SYQVDVFCNSEDGDYALTPKSSYAFGR-----SASSLDVNASGKMMVHGS	959
QY	1135	-ELDIQHNQIKTS-----EIRHRCFGGOYQITALPNNGNHLQYHDOQODSCASLSNS	1188
Db	960	IFLDI--STGRSSDVSIDNVKQCFKPEAEV-----LATSR	993
QY	1189	KTQSSANT-----TQSNTPESATYMKLIDLERERTQONKASPTDAAK--VITYSIMA	1238
Db	994	ESKNIAITVVEAARSWTDYPEKSDATIDENSTGTPRLEGRLAIVSDGRGAVISTVEDA	1053
QY	1239	ENGELILASINMESEVVRV-----ALTYTKQAPELLQTSILTN	1276
Db	1054	QEG-----NPLVELSSVKTIPHTSNGSPILPOLINQISVSEVTLDKRELAQOLVTA	1108
QY	1277	LGICIGKGNCELPNKGFPSIRMRMLNIIIEAHTDVYAGLHVLSKIMRSNK--MRHNMH	1334
Db	1109	-----NDNSIMTKYFENOILTTLILEVDLDDSDSISRELSTLIVAEMLHNOXDPMES	1158
QY	1335	FLELILKTIQCYQH-----SKELARDIDSMIPR-----IAPSLPDLSTININVPAT	1383
Db	1159	-TEIVLEKLIVTKQVAKVNSNAQOCINNVYAKIDPRCLAVIYPLVLVSDEKMLVCT	1217
QY	1384	GEFPTNLCAIKILLEVEHSGEITDALDIYFPYPLARSADDTQSMKRAKAVFCIKLYF	1443
Db	1218	N-----CLTLYVGRLSR-----ELMTQOLPBLFALYPAFNQNSPDVAKTYVFC	1261
QY	1444	VUGEKVKPKULSVLNPKSVRLNLYIEKQRCNISG	1478

[illegible]

PRIOR FILING DATE: 2002-03-05
 PRIOR APPLICATION NUMBER: GB 0105401.4
 PRIOR FILING DATE: 2001-05-03
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: Patent In Ver. 3.2
 SEQ ID NO 4
 LENGTH: 2065
 TYPE: PRT
 ORGANISM: Xenopus laevis
 US-10-469-780-4

Query Match 4.7%; Score 358.5; DB 4; Length 2065;
 Best Local Similarity 19.3%; Pred. No. 2,7e-14;
 Matches 311; Conservative 266; Mismatches 588; Indels 449; Gaps 70;

13 IQGPKADMKVVKQALDEL-VTFLESDPTNSIVCTDGMFLIDGMLPMLTGSHPFLAKQSLK 71
 602 MQQDSSNMKERLASMEEFOKTVESMERNDIPCOALVKML-AKPKGKETNFQVMQMKH 660
 72 AFSELIKRLGSDFNAYATATVLPVVIDRLGDSRDTREKAQLLRDLMHRLVPPQALIDK 131
 661 IVA-LIAOKGNFSTKSAVALDGLNDKVDYK-CGNAKAEALSGIAEACTLPMTA--EQ 715
 132 LAISCFHKAQAKYREELQITIVAAHVEYTOQLSVRYIIPVYCALLGDPVANNREAAIQ 131
 716 VSLAFQKPKKQKQSELTNMLSNAIKEFGTGIVKAFISNVKTALATNPALRTSAITL 775
 192 LVEIYKAVGDRLEPRMDVPAKSLAMEOKFPDQKQEGLLPSPALKNTNGVGLDE 251
 776 LGWVLYMGAPLKMFE--EKKRA-LLSQIDAEEKAK--GQTPPVYSIRSK-HGSRDE 829
 252 ADNIGLERPRTMIKRLHSAVSSSLRPKPNVDVTGACAVTWESFEVVPOLNIF 311
 830 GE-----EGEQDEDPADAVT-----DILPRTDI- 853
 312 HAQMDIYQOVLLITSDKNADEKVDALKIRALLILSYHPQPVANVQLKELSLFV 371
 854 ----SDKISSDLVSKIEDKN--WKIRKEGLDEVTAII--NEAKFIQPSIELPALK 902
 372 DILKEELRSQVIREACITIAVMSKTLNKLDAFCMSLIEHLINL-----IONSAKIVA 424
 903 GRILAND--KKILVQOTITLQOULST-----AMKHNIKQHKVGLMFIITVLGDSKANVR 954
 425 SASITALKYIITKYTHAPKLIKITYDTLNOSKSKDIRSTLCEMLVLLFEWQTKALENNAT 484
 955 AALGLTKSWVDQGMKWLE--GEDLSEELKKNPFLROELL-----GMLAEKLPEMRT 1007
 485 VLBDT-----LKKSIDADCDARRHSRYAVMARRRFPPELADQIYGTLDIAAORALERE 538
 1008 VPEDLQLCVEFYLVNCLIEDRNGDVAKKQAEALPIFMHMI-----GFEK- 1049
 539 REGGGGGGTGTGCTAPETRTVRIRIGRTPTLOKTPPSMRSISAVDTAAORAKVRAQY 598
 1050 -----MSKATSKILPKASK-DQVVALLEKAKA-- 1074
 599 TLVSROKRLPGPNNSNQASMTGAASGL-PRPRLNSGCTPA-----TTPGVT 649
 1075 ---SMPAKPAGPPKASSKQPPAVAAQASAPPAASDSGSGSTDYKPPKTKPTGTQAS 1131
 650 RPRGRAGVSOQPSRSTSPSTKRDQYGGIYRRATGAIPIKAGAGIRSTASSET 709
 1132 KAKTQ---SVSSEGTSLNPSN-----TSITPSPKATSLSKAKPAKQTL 1172
 710 PTRSGGGLMKRSMYSTGASGRTRPERNNPVPAPAPARLLAQSRAREHTLGVGDGQPDYV 769
 1173 PGK-----KAPSKPNKDEB-----DMSGIYI 1195
 770 SGDYMRSGMRGKMLGRDESDDIDEBASVCSERSFDSSYTRGKSNYSLSGSHTRLD 829
 1196 I-----VNGKEQVRKDE-----KALKV-----LK 1215
 830 W--STORAPF-DLLEITIIQFCASTHMSERK-----DGLISLTOYLADGKELTQQQ 876

1216 MNFTPRDEYITQNTQMSPCIA-RWLODELPHADPQRIKGLAVMTHELSBKGVISC 1274
 877 LKCVLDMERKEMDTHTKYSLPLDTVELLIVHANE-----TSRNGSSCLTRLPFKLG- 931
 1275 LDVLKMFTRPFDNTSVLMKCLEYLLKFLYMLQOEYHILTEMGG-TSFLYLMKVG 1333
 932 -TDLL-NSMHSKIKWTQVNHVEYPTQOLKSLFPIISDSTQPTTKTIALRFLTDA 989
 1334 PDIVAKDVRALLTKRCOV---YPA---SKMFNVMEGTYSKSKQBAECLEELGCLV 1385
 990 NTYCKSSDPFPOSOACERT---VLKLAQIADQKSMELRSQARGLVALYNINTPOMTL 1046
 1386 ESY-----GANVCQPTPAKALKETAIHIGDDDTYRNALNTIYVYVHGGQVRK 1436
 1047 LADLPKYQDSARSCISHMRQOSCSNGANSPPSSPLSSSPKPLQSPVGPASIQ 1106
 1437 LIGNL---SEKQMSLEIRIKR-----AGKQAAAPAKQVEEKPQVQSAN--ASIL 1484
 1107 SHHQLSTISIPRER-----QSVQOELLFSELD-----IQHNI 1142
 1485 RAAPEDMSKLNQARNMGHTPEPSHVRE--FOLDLEIENDNGTVACEMPALYQHL 1542
 1143 OKTSE-----EIRHCFGGQYQOTALAPNGFNGHLQYHDQGOQDSCASLSNSTOSSAN 1195
 1543 DEIFEPVLIPEKIR-----AVSPHFDMH-----SNTASTIN 1575
 1196 -----TQSNTPSAPATRLDLEBERTQNAKSEPTDAKIT--VSINMAENGEL--- 1243
 1576 FYISQVAVDINASIQALQIDEVLQEDKAEAMSGHIDQFLIAFTMQRLAYNTMADE 1635
 1244 -----ILASNL-----MESEVVRVALTITKQDPVELLOSLNLNIGICIGKNG 1286
 1636 RLKDQDIVELYSCTIGNMISLFPOMESLAREASTGVAKD-----LHGLSLMLDARIED 1690
 1287 ELPNKHFRIEMMLNILE-AEHTDVIAGLVLSKIMSNNKRRHMMHFLBILKIIQ 1345
 1691 EGGQYVRSVNLVYKLEKSDQNTNISALLMLDQSLATASSPFSSELVMMKCLRMIR 1750
 1346 CYQHS-----KEALRDISMTPIRIAPSLPL-----DLSINIVPIAIGEFPTNCAIK 1394
 1751 LPEAINNLNLRILLDINHFM-RVLPKEKELQHKSEMWRTLKTLHT-----LCKLK 1803
 1395 -----ILLEVEHHSSETTDHLDIVFPNLABSADDTQSVWRK--AAVFCIVKYFVL 1445
 1804 GPKINDHLSMIENKHESEL-EAHLLRV--MKHSIDRTSKGDKETEKASCTI----- 1852
 1446 GEEKVK-----PKLSVLPNSKYRLINVYIEKORNCISGSGSTKNSS 1487
 1853 -EDKVKANVSDFLAEMPKKIGSKENTKEGLAEIYKKKYSDADIKPFLKSS 1905

RESULT 8
 US-10-408-765A-263
 Sequence 263, Application US/10408765A
 Publication No. US20040101874A1
 GENERAL INFORMATION:
 APPLICANT: Ghosh, Soumitra S.
 APPLICANT: Fahy, Bojin D.
 APPLICANT: Zhang, Bing
 APPLICANT: Gibson, Bradford W.
 APPLICANT: Taylor, Steven W.
 APPLICANT: Glenn, Gary M.
 APPLICANT: Watnick, Dale E.
 TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
 FILE REFERENCE: 660088, 465
 CURRENT APPLICATION NUMBER: US/10/408, 765A
 CURRENT FILING DATE: 2003-04-04
 NUMBER OF SEQ ID NOS: 3077
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 263
 LENGTH: 1972
 TYPE: PRT

ORGANISM: Homo sapiens
US-10-408-765A-263

Query Match 4.3%; Score 328; DB 4; Length 1972;
Best Local Similarity 18.7%; Pred. No. 2.9e-12;
Matches 278; Conservative 253; Mismatches 543; Indels 416; Gaps 68;

```

QY 61 SHEKIAQKSLAEFSELIKRLGSDFNATATVLPVHVIDRLGSDRDYREKAQILLRLDLMH 120
DB 653 TNFOVMQMKLHIVA-LIAQKGNFSKTSQAQVLDGLVDKIGDVK--CGNNAKEMAMTAIAEA 709
QY 121 RVLPPQALIDKLATSCPFKNAKAREEFLQTIYNALHEVTOQLSRYVIPPYCALLGDP 180
DB 710 CMLPMTA--EQVVSMAFSQKNPKQSEETLNMWLSNAIKFEGFSGLNKAFISNVKTALAAAT 767
QY 181 TVNVEAAIQTLVEIYKHVGRRLRPDLRRM--DDVPASKLAMEQKFDYQOEGILLPSA 238
DB 768 NPAVTRTAITLGLVMVLYVG---PSLRMFEDKRA--LLSQIDAFFE--KMGQSPAP 820
QY 239 LKNTNGNGV-GLDEADNIGLERPRTMIKRLPHSAVSSSLRPKNVDYTGAGAVTMES 297
DB 821 TGISKHSSTGTDEGD-----GDEPDGSDNDV-----849
QY 298 FESSFEVQOLNIFAKMDMDIYKQVLIISDKNADMEKVDLKKIRALLISHTQO 357
DB 850 ----DLRPTBL-----SDKITSELVSKIGDN--WKIRKEGLDVAIGI-----NDAK 892
QY 358 FVAVQKELSLFVDILKEELRSQVIREACTIAYMSKTLRNKLDAFCWSILEHLINLQ 417
DB 893 FIOPNIGELPTA---LKGRINDS-----NKI-----LVQOTLNTILO 925
QY 418 NSAKVIASATLAKYIIKYTHAPKLLKIYTDPLNOSKSKDIRSTLCMLVILLFEEMQTK 477
DB 926 Q---LAVAMGPNIQOHVKNLGIPI-ITVLGDSKNVRAAALATV-----965
QY 478 ALERNAIVRLDTLKSGIGDADCDARRSHRAYAFRRHPELADQIYGTLDIAQAALER 537
DB 966 ----NMAAQOTGKEMLEGEDELSE-----LKKENPFRLQELMGW--AEK-----1005
QY 538 EREGGGGGTGTGTGAPETRRVSRIGRTPTLQKPTSMRSISAVDTAAQOR--ALV 594
DB 1006 -----LPTLRST-----PTDLILCVPLHYSCLEBRNDVAK 1036
QY 595 RAQYTL-----YGRORRPLG--PNSNQASMTGAASGSLPPRLNNSGGTPTAT 643
DB 1037 KAQDALPFRPMHLGYEKMAKATOKLKPSTKQD-----VLAMLEKAKVNMPPAKPAP--1086
QY 644 PGSVTPRPGRACVSSQSPGSRST---SPSTKLBDQYGGIGNYRGATGALPKKASGIP 699
DB 1087 PTKATSKPMGGSAPAFKOPASAPAEDCISSSTEPKPD-----PKKA-----1127
QY 700 RSTASRETSPTRSSGGLMKRSMYSTGAGSRRTPENNNPVPSAPARLLAQSREAHITLG 759
DB 1128 -----KAPGLSSKAKSAQG--KKMPSKTS-----LKEDF-----1154
QY 760 VGDGQPDYV---SGDYMSGGWRMGRKLM-----GRDE-SPDDISEASSVCS--BR 805
DB 1155 --DKSGPITIVPNGKEQRMKDEK-GLAKYLKMNFTPRDEVIQOLTKQSSCAVAKWLODE 1211
QY 806 SFDSSTYRKNSKNSYLSGSHTRLDMSTQAPAPDDITITIIQFCASHTWSEKQGLISLTOY 865
DB 1212 MFSHSDFOHNNKA-----LAVMYD-----HLESEKGVIGCC--1241
QY 866 LADGKELTGOQKCVLDMFRKMPMDHTTKYVSLFLTVTELLIVHANE---TSRNGSSC 922
DB 1242 -----LDLILKMLTLFFDTNITSYLMKALYKLLFLLISEEYHLTENASSF 1290
QY 923 LTRLFNKLQ--TDLNLSMHSKIMKTLQVVEHYEFTLOLKELEFRIISDSQTPTTKRIA 980
DB 1291 IPIVLVYKGPVKVIRKIDRALINRMCLVY-----PASKMPFIMEGKSKNSKQRAE 1343
QY 981 IIRFLTDLANTYKSSDPSDSQACERT--VLKLAQLAADOKSMELRSQARSCVALY 1037

```

```

DB 1344 CLEELGCLVESY-----GNANVCQPTPGKALKEIAVHIGDRBNVANAALNTIVTY 1394
QY 1038 NINTPQMTLLADLPKRVYODSARSCIHSHMRQOSCNNGANSFSSSPSSPKFLOSP 1097
DB 1395 NVHGQVFLIGNLNTL-----SEKMSMLEERIKR-----SARPSAALIKQVEKQRAQ 1443
QY 1098 SV-----GPFASLQSHHQLSISTSPRSRQSVQELLFSELDIGNICT--1145
DB 1444 NISSNANMLRKQPADMSKLNQARSMGHPBAQ--MYRE--FOLDDEINDNGTVRC 1500
QY 1146 --SEELRHQFGQYDTALAPNGFNGLQYH--DQGQODSCASLSSNKTOSANTQSNTP 1202
DB 1501 EMPELYQKLDLDFEPLVILPEPKIRAVSBPHDMSNTASTNFIISQVASGDINTSQA 1560
QY 1203 ESATRLNLERERTTONAKSPPTDAKYITVSI-----NMAANGELILASNMESEVAV 1257
DB 1561 LTQLFOISLAEASTGVKDLMHG--LITLMLDSRIEDLEGQYIRSVNLLVYVLE- 1617
QY 1258 ALTLTKQDQ-----VELLOTSLTNLGICIKGNCBELPKHFRSITRMMLNILEAHTDV 1311
DB 1618 ----KSDQTNILSALLVLLQDSL--LATASSPFSSELVWKCLMRVRLLPDTINSINDR 1671
QY 1312 VIAGLHVLSKIMRSNKMHNMMHFLLELLIKTIQCYOHSKEALRPIDSMI-----PR 1363
DB 1672 ILDDHIFMKYPPKKLK-----QC--KSEFIRLTKLTLHTLCKLQKPK 1714
QY 1364 IAPSLPLDLSINIYNVPIATGEPPTNLCAKILLEVTEHSGEITDAHLDIVPMLARSA 1423
DB 1715 ILDHLTMIDNKN-----ESELHAHLCRM-----KHSMDQGSKSD--KETAKA 1757
QY 1424 D--DTQSMVRKAAYFCIVKLYFVLG--BEKVPKLSVLANPSKRLNLNYYE 1470
DB 1758 SRIDAKSSKAKVNDP-LAEIFKIGSKENTYKGLAEIYKKKYSADADIE 1806

RESULT 9
US-10-469-780-6
; Sequence 6, Application US/10469780
; Publication No. US20040147724A1
; GENERAL INFORMATION:
; APPLICANT: FAGAN, RICHARD JOSEPH
; TITLE OF INVENTION: NUCLEAR HORMONE RECEPTOR LIGAND BINDING DOMAIN
; FILE REFERENCE: 015959-00003
; CURRENT APPLICATION NUMBER: US/10/469,780
; PRIOR FILING DATE: 2004-02-04
; PRIOR APPLICATION NUMBER: PCT/GB02/00941
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: GB 0105401.4
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 6
; LENGTH: 1972
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-469-780-6

Query Match 4.3%; Score 328; DB 4; Length 1972;
Best Local Similarity 18.7%; Pred. No. 2.9e-12;
Matches 278; Conservative 253; Mismatches 543; Indels 416; Gaps 68;

```

QY 239 LKNTNGCV-GLEADNIGLERPRTMRKPLHSAVSSSLRKPKNVDVTGDAVATMES 297
 Db 821 TRGISKHSSTGDEGED-----GDEPDGSDNV-----849
 QY 298 FESFEVVPQAINFIKAKMDDIYKOVVLIISDKNADWEKVDALKIRALLILSYHTPOQ 357
 Db 850 -----DLPRTEL-----SDKITSELVSKIGDN--WKIRKGEDEVAGII-----NDAK 892
 QY 358 FVAVOLKELSLFDVILKEELRSQVIRACITIAVMSKTLRNKLDAFCMSILIEHLINLIQ 417
 Db 893 FIQENIGELPTA-----LGRLANDS-----NKI-----LVQGTINILIQ 925
 QY 418 NSAKVIASASTIAKIIKYTHAPLKIYDTLNOSKXDIRSTLCMLVLLFEEMOTK 477
 Db 926 Q-----LAVAMGNPKOHVKNLGIP--IITVLGDSKNVRAALATV-----965
 QY 478 ALERNATVRLTKKSIDGADCDARHRSRYVMAFRHFPPELADQIYGTLDIAAQRALER 537
 Db 966 -----NAMAEOGMKEWLEGEDELSEB-----LKENPFLROELGWL-----AEK-----1005
 QY 538 EREGGGGGTGTGTAPETRTVSRIGRTPTLOKPPMSRSISAVDTAAOR-----AKV 594
 Db 1006 -----LPTLRST-----PTDLILCVPHLYSCLEDNNGDVVK 1036
 QY 595 RAQVYL-----YSROKPLG--PNNNSQASMTGAAGSLPRPLNSNGGTPATT 643
 Db 1037 KAQDALPFFMHLGVEKAKATGKLIKPSKDQ-----VLAMEKAKVMNAPAPAP---1086
 QY 644 PGSTVPRGRAGVYQSOPGSRST--SPSTKLADYOGIGYKATGAIIPKASGIP 699
 Db 1087 PTKATSKRMGSAKAPQAPAPADCISSSTEEKPD-----PKKA-----1127
 QY 700 RSTASSSETSPTRSGGGLMKRSYSTGASRTRPENNPVPAAPARLLAQRAREHTLG 759
 Db 1128 -----KAPLSSAKAQAQ-----KMSKTS-----LKDE-----1154
 QY 760 VGDDGQPDVY--SGDYWRSGMGRKLM-----GRDE--SDDIDBASVCS-----ER 805
 Db 1155 --DKSGFIIVPNQKEORMDEK--GLKVLKMNFTTPREYIEQLOKQMSCVAKWLQDE 1211
 QY 806 SFDSYTRGKNSVSLSGSHRLDMSTORAFDDIETIIQCASTHNSERDGLISLTQY 865
 Db 1212 MFHSDPFHNNKA-----LAVAVD-----HESEKEGVIGC-----1241
 QY 866 LADKELTQOOLKCVLDFRKMFMHTHKVYSLFDVTTELIVHANE--TSNNGSSC 922
 Db 1242 -----LDLILKMLTLRFDDINTSVLMKALEYIKLFTLLSEEBEYHLENASSF 1290
 QY 923 LTRLEFNKLG--TDLLNSMHSKIMKTLQVHEHYFTOLOKELFR1ISDSTOTPTTKTRIA 980
 Db 1291 IPYLVVAKGEPKQVIRKQVRAILNRMCLVY-----PASKMFPFIMEGTSKNSKQRAE 1343
 QY 981 ILRLTLTLANTYCCSSDPSPQOACERT--VLKLAOLADQSMELRSARSCVALY 1037
 Db 1344 CLEELGCLVESY-----GNNVCOPTPKALKEIAVHIGDRDNVRAALNTIYVY 1394
 QY 1038 NLNTPOMTULLADLPKYVQDSARCSISHMRQOSGNSGANSPPSSPLSSSPKPOSP 1097
 Db 1395 NVHGDQVFKLGNL-----SECDMSMLERIK-----SAKRSAAPIKOVEKFORAQ 1443
 QY 1098 SV-----GPRASIOSHHOLISISTSPRSROSSVEOELFRSSELDIOHNIQKT--1145
 Db 1444 NISSANAMLRKGPADEMSSKLNQARSMSGHEBAQ--WVRE--FOLDIDEINDNGVIRC 1500
 QY 1146 --SEIHCFCGGOVOTALPNGFNHLOYH--DOGOQOSCASLSSNSTQSSANTQNTP 1202
 Db 1501 EMPELVQHKLDITEPVLIPEPKIRAVSPHFDNHSNTASTINFIISQVAGSDINTYIOA 1560
 QY 1203 ESATMRDLNERERTQNAKSPTDKAYTVSI-----NMAENGEILLASLSESEVVR 1257
 Db 1561 LTOFQIESLARKEASTVGLKOLMNG--LITLMDLSRIEDLEEGOVIRSVNLLVVKYLE-1617

QY 1258 ALTLTQDP-----VELIQTSLTNIGICIKGENCELPNKHFRSINRMLNIEAERTV 1311
 Db 1618 -----KSDQTNILSALLVLQDSL--LATASSPKFSELVKKLCMRWRLLPDTINSINLDR 1671
 QY 1312 VIAGLHVLSKINSNKGRRNMMHFFLELILKIIIOCYQHSKEALRDISMI-----PR 1363
 Db 1672 ILDDIHFPMKVPEKELK-----QC--KSEFPRTKITLILTLCKLKGPK 1714
 QY 1364 IAPSLPDISINIVNEVIATGEPPTNLCAIKILLEVEYHNGSEITDAHLDIYEPNLARSA 1423
 Db 1715 ILDLHLMINKN-----ESELHAHLCRM-----KXSMOQTSKSD--KETAKGA 1757
 QY 1424 D--DTOSMVRKAAPFCIVKLYVLG--BEVYKPLASTLANSKRLANVYE 1470
 Db 1758 SRIDAKSSAKVNDP-LAEIIFKXIGSKENTYKGLAEIYEVKKYQSPADIE 1806
 RESULT 10
 US-10-469-780-2
 ; Sequence 2, Application US/10469780
 ; Publication No. US20040147724A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FAGAN, RICHARD JOSEPH
 ; APPLICANT: PHELPS, CHRISTOPHER BENJAMIN
 ; TITLE OF INVENTION: NUCLEAR HORMONE RECEPTOR LIGAND BINDING DOMAIN
 ; FILE REFERENCE: 015959-00003
 ; CURRENT APPLICATION NUMBER: US/10/469,780
 ; CURRENT FILING DATE: 2004-02-04
 ; PRIOR APPLICATION NUMBER: PCT/GB02/00941
 ; PRIOR FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: GB 0105401.4
 ; PRIOR FILING DATE: 2001-05-03
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 3.2
 ; SEQ ID NO 2
 ; LENGTH: 2038
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-469-780-2
 Query Match 4.1%; Score 314.5; DB 4; Length 2038;
 Best Local Similarity 18.6%; Pred. No. 2.4e-11;
 Matches 271; Conservative 238; Mismatches 517; Indels 433; Gaps 65;
 QY 61 SHFKIAQSLSEAFSELIKRLGSDPNAYTATVLPVHVIDRLGDSRDTVREKAQLRLDMEX 120
 Db 659 TNQVWQMLHIVA-LIAQKGNFSKSAQVLDGLVDKIGDYK--CGNNAKEMTAIAEA 715
 QY 121 RVLPQALIDKATSCFKRKAQVREELQTVNALHEVGTQOLSRYVYIPVYCALLGDP 180
 Db 716 CMLPMTA--EYVVSMAFSQKNPNQOSETLNMLSNAIKERFGSLNKAFLSNVKTALAAT 773
 QY 181 TVNVRERAIOTLVEIKHVGDRLPDLRRM--DDVPAKSLAMJEKFDQVQKQGLLPSA 238
 Db 774 NPVARTALITLGLWMLYVG--PSLRMFEDDEKA-LISQIDAEFE--KMGQSGPPAP 826
 QY 239 LKNTNGCV-GLEADNIGLERPRTMRKPLHSAVSSSLRKPKNVDVTGDAVATMES 297
 Db 827 TRGISKHSSTGDEGED-----GDEPDGSDNV-----855
 QY 298 FESFEVVPQAINFIKAKMDDIYKOVVLIISDKNADWEKVDALKIRALLILSYHTPOQ 357
 Db 856 -----DLPRTEL-----SDKITSELVSKIGDN--WKIRKGEDEVAGII-----NDAK 898
 QY 358 FVAVOLKELSLFDVILKEELRSQVIRACITIAVMSKTLRNKLDAFCMSILIEHLINLIQ 417
 Db 899 FIQENIGELPTA-----LGRLANDS-----NKI-----LVQGTINILIQ 931
 QY 418 NSAKVIASASTIAKIIKYTHAPLKIYDTLNOSKXDIRSTLCMLVLLFEEMOTK 477
 Db 932 Q-----LAVAMGNPKOHVKNLGIP--IITVLGDSKNVRAALATV-----971
 QY 478 ALERNATVRLTKKSIDGADCDARHRSRYVMAFRHFPPELADQIYGTLDIAAQRALER 537

```

Db      972  ---NAAEQTGMKEWEGEDLSE-----LKKENPFLROELIGWL-----AEK----- 1011
Qy      538  EREGGGGGTGTGTGTATPETRRATVRIGRPTGLQKTPPMRSISAVDTAAQR---AKV 594
Db      1012  -----LPTLRST-----FTDLICVPHLYSCLEDRNGVRK 1042
Qy      595  RAQYTL-----YSRQRKPLG---FNNSNQASMTGAAASGLPRRLNSNSGTPATT 643
Db      1043  KADALPFFMHLGYEKAKATGKLTKPTSKDQ-----VLAMLEKAKVMPAPAP--- 1092
Qy      644  PGSVTPRRRAGVSGSQPSGRST---SPSTKRDQYGGIGNYRATGATAPKASGIP 699
Db      1093  PTKATSPKMGSAAPAKQOPASAPAEDCISSSTEKPD-----PKKA----- 1133
Qy      700  RSTASSRSTPTRSGGGLMKRSMYSTGASRRTPERNRPAPAPARLLASRAEHTLG 759
Db      1134  -----KAPLSSKAKSAQG---KMPSKTS-----LKEDE----- 1160
Qy      760  VGDDGQPDYV---SGDYRSQGMKMRKLM-----GRDE-SDIDSEASSVCS---ER 805
Db      1161  --DKSGIFIVPENGKEQRMKDEK-GLKVLKMNFTTRDEYIEGLKQMSQVAKWLODE 1217
Qy      806  SFPSSYTRGKSNYSLSGSHTRLDWSTQRAPPDIEITIQFCASHTWSEKQGLISTQY 865
Db      1218  MFHSDPQHNKA-----LAVMVD-----HLESEKEGVIGC--- 1247
Qy      866  LADGKELTQOQLKCVLDFMRKMFMDHTKYSLFLDVTELLVHANE---TSRNGSSC 922
Db      1248  -----LDLILKMLTLRFPTNTSVLMKALEYKLLKLLISEEYHLTENESSF 1296
Qy      923  LTRLFNKLG---TDLNSMHSKIWKTLQVNEHYEPTQLKELFRLLISDSTQPTTKTRIA 980
Db      1297  IPLYLVKVGPKQVIRQDVAILNRMCVY-----PASKMPEFMEGTRKSNKQRAE 1349
Qy      981  ILFLFDLNTATYCKSSPFSQOACERT---VLKLAQLAADQSMELRQASCLVALY 1037
Db      1350  CLEBELGLVSEY-----GMNVCOPTPKALKELAVHIGRDVNRANAANTITVY 1400
Qy      1038  NLATPQWTLTLLADLPKYVODSARSCHSHMRQSQSCNSGNSPSSPLSSSPKPLOSP 1097
Db      1401  NVHGDOVFKLIGNU---SEKDSMLEERIKR-----SAKRPSAAPIQOVEERKQRAQ 1449
Qy      1098  SV-----GPFASLQSHHQLSISSTSPRSROSSVEQELLFSSELDIOHNIQKT--- 1145
Db      1450  NISSNANMLKQGAEDMSKLNQARSMSGHPBAQ--WVRE--FQDLDEIENDNGVRC 1506
Qy      1146  --SEIRHCGGQYQATALAPNGFNGLQYHDQGOQSCASLSSNSKTQSSANTQSTPE 1203
Db      1507  EMPELYQHKLDDIFEPVLIEPPIKRAVSPHFDMSHTASTINFIISQVASG--DINTSI 1564
Qy      1204  SATMRDLNERERTTQNAKSPPTDQAKIYTSINMAEGELLILASNMESVAVALLTK 1263
Db      1565  QALTQIDEVLRQDKAEAMSGHIDQFLIATFMQ---RLIYTHMADEK-----LEK 1613
Qy      1264  DQPEVL-----LQSLTNLQICIGCIGCCELPNKHFS 1295
Db      1614  DEIKXLYSCIGNMISIFQIESLAREASTGVLDKMLMGLTLMDSRIDELEGQGVIRS 1673
Qy      1296  IMEMLANILE-AEHTDVVIAGLVLSKIMRSNQRHNMHFLILIKITIQCYOHSKAL 1354
Db      1674  VNLVVVLEKSDQNTLISALLVLODSLILATASSPK---FSEIWM-----KCL 1719
Qy      1355  RQDSMIRIAPRLPD---LSININPVY-----ATGEFPL-----NLCAIKIL 1396
Db      1720  WRVVRLLPRTINSINDRILDLHIEMKVPPEKELKQCKSEFIRTLKTLHLTLCKLK-- 1777
Qy      1397  LEVTEHNGSEITDAHLDIV 1415
Db      1778  -----GPKILD-HLIMI 1788

```

RESULT 11

```

US-10-307-928A-16
; Sequence 16, Application US/10307928A
; Publication No. US20030229016A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, John P.
; APPLICANT: Anderson, David W.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Caterton, Elina
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xieojia (Sasha)
; APPLICANT: Li, Weizhen
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: Patrutajan, Meera
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
; FILE REFERENCE: 24102-502D
; CURRENT APPLICATION NUMBER: US/10/307,928A
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: 60/341,477
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/341,540
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/342,592
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/344,903
; PRIOR FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/373,288
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/380,981
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 60/381,495
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/383,744
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: 60/384,024
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/401,788
; PRIOR FILING DATE: 2002-08-07
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: CuroSeqdlet version 0.1
; SEQ ID NO 16
; LENGTH: 2040
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-307-928A-16

Query Match      4.0%; Score 304.5; DB 4; Length 2040;
Best Local Similarity 18.5%; Pred. No. 1.2e-10;
Matches 271; Conservative 239; Mismatches 516; Indels 441; Gaps 66;

Qy      61  SHEKIAQKLEAFSELIKLGSDFNAYTATVLPVVIDRLGDSRDYVERKAQLILRLMEH 120
Db      653  TNFQVWQMTLIVA-LIAQGNFSTSAQVVLGVLGDKIGDYK--CGDNAKEAMTAIAEA 709
Qy      121  RVLPQALIDKLATSGFRKNAKAREEPLOTIVNALHEGTQQLSVRYVTPPYCALLGP 180
Db      710  CMLPMTA--EQVNAAFSQKPKNSQSETLNLMSNAIKERFGSLNKAFTISNVKTALAA 767
Qy      181  TVNVEEAATQTLVEYIKHVGDRLLRPDLRPM--DVPASGLAMLEQKFDQVQKGLLPSA 238
Db      768  NFAVRTAATTLGWTLYVG-----PSLRNFFEDKRA-LISQIDAEFE--KNGQSPAP 820
Qy      239  LKNTNGNGV-GLDEADNIGLRERPTRMIRKPLHSAVSSSLRPKPNVNDVTGAGAVTMS 297

```

```

Db      : : : : :
821 TRGISKSTSGTDEED-----GDEPDGSGNVV----- 849

Qy      : : : : :
298 FESFEVVPQINIFHAKMDDIYKOVAVIISDKNADMEKRYDALAKIRALLILSYHTPOQ 357
: : : : :
Db      : : : : :
850 ----DLIRREI-----SDKITSLSVKIGDKN--WKIRKEGLDEVAGII-----NNAK 892
: : : : :
Qy      : : : : :
358 FVAVOLKELSLSPVDILKEELRSQVIREACTTIAVMSKTLNKKDAPCWSLLEHLINLIQ 417
: : : : :
Db      : : : : :
893 FIQNIIGELPFA--LKGRLNDS-----NKI-----LVQOTLINLIQ 925
: : : : :
Qy      : : : : :
418 NSAVIVASASTIALKVIITKHAPKLIYTDITLNOSSKOIRSTLCMLVLEFEWYTK 477
: : : : :
Db      : : : : :
926 Q--LAVAMPBNIKOHVKNLGI-P-IITVLGDSKNVNAALATV----- 965
: : : : :
Qy      : : : : :
478 ALERNAFVLRDYLKKSIGDADCDARHRSYAWAFRRFPPELADQIYGLDIAAORALER 537
: : : : :
Db      : : : : :
966 ----NAMEQOTGMKEMLEGEDELSE-----LKKNPFLRQELLGWL--AEK----- 1005
: : : : :
Qy      : : : : :
538 EREGGGGGGTGTGTGTAPETRRYVSIRIGRTPTGLQKPTPSMRSISAVDTAAQR--AKV 594
: : : : :
Db      : : : : :
1006 ----LPTLRST-----PTDLICVPHLYSCLEDNRGDIYRK 1036
: : : : :
Qy      : : : : :
595 RAQYTL-----YSRQKPLG--PNNNOASMTGALASGSLPRRLNNSGTPATT 643
: : : : :
Db      : : : : :
1037 KAQDALPFPMMHLGEXKAKATGKLKPTSKO-----VLAMEKAVNMPAPAP-- 1086
: : : : :
Qy      : : : : :
644 PGSTPRGRAGVSOQSGSRST--SPSTKLDYGGIGNYRGATGAI PKKASGIP 699
: : : : :
Db      : : : : :
1087 PTKATSKMGSAIPAKFQAPAPADCCISSSTERPD-----PKKA----- 1127
: : : : :
Qy      : : : : :
700 RSTASSRETSPTSGGGLMKRSMYSTGAGSRTPRRNNPVPAPARLLAOSREAHETLG 759
: : : : :
Db      : : : : :
1128 ----KAPGLSSKXSAQ--KMPKST-----LKEDE----- 1154
: : : : :
Qy      : : : : :
760 VGDDGQPRVY--SGDYRSGGMWGRKLM-----GDE--SDDISASVCS-----ER 805
: : : : :
Db      : : : : :
1155 --DKSGPIFIVVPGKEQRMDEK--GLVYLMKNFTTPDEYIEQIKTMSCVAKMLODE 1211
: : : : :
Qy      : : : : :
806 SPDSYTFGRKSNVSLSGSHTRLDWSTQARFDDIETIIQFCASHTWSEKDGILSLTOY 865
: : : : :
Db      : : : : :
1212 MFHSDFOHNNKA-----LAVVND-----HLESEKGVIGC-- 1241
: : : : :
Qy      : : : : :
866 LADGKELTQOOLKCVLDMFRKMFMDTHTKVYSLFLDYTELILVHANE--TSRNGSSC 922
: : : : :
Db      : : : : :
1242 ----LDLILKWLTRFPEDTNSVLMKALEVILKLFLLSEEEYHLTEHASSF 1290
: : : : :
Qy      : : : : :
923 LTRLPNKLUG--TDLLNSHNSKIMKTLOVYVHEFYQLOLAKELFRILISSTOTPTTKRIA 980
: : : : :
Db      : : : : :
1291 IPYLVVKKGEKPDVIRKDVRAILNNMCVY-----PASKMFPTIMGTSKSKSKORAE 1343
: : : : :
Qy      : : : : :
981 ILRFLTDLANTYCKSSDFPSDOSQACERT--VLKLAOLADOKSMELRSOARSGCLVLY 1037
: : : : :
Db      : : : : :
1344 CLEELGCLVEY-----GANVCQPTBGKALKEIVNHGDRNAVNAALNTIYVY 1394
: : : : :
Qy      : : : : :
1038 NLANTPQMTLLADLPKYVQDSARSCISHMRROQSCNSGANSPPSSSPKXPLQSP 1097
: : : : :
Db      : : : : :
1395 NVHGQVFKLIGNL--SEKDMSWLEBRIKRS-----AKRPAADIKOVEEKPQAO 1443
: : : : :
Qy      : : : : :
1098 SV-----GPRASLOS-----HHQOLISSTSPRSROSSVEBELLFSSBLDIOH 1140
: : : : :
Db      : : : : :
1444 NISSNANMLRKGPADMSSKUKIMYRTYRIQARNSGHPAAQ--VVRRE--FOJLDLEIE 1500
: : : : :
Qy      : : : : :
1441 NIQKT-----SEIRHCGGOYOTALADNGFNGLQYVDOQOQSCASLSNSKTOSSAN 1195
: : : : :
Db      : : : : :
1501 NDNGVRCMEPEIVQHKLDIIFEPVLIPEPKIRANSPHFDMDHSTASTINFIISQVANS 1560
: : : : :
Qy      : : : : :
1196 TTQSTPESATMRLDNLERERTTONAKSPETDQAKYIVYSIMANGELILASNLMESEV 1255
: : : : :
Db      : : : : :
1561 --DINTSIOALTQIDEVLRQEDKAEAMSGHIDQFIATFMQI--RLIYVTHNADEK-- 1612
: : : : :
Qy      : : : : :
1256 RVALTLTDOQVEL-----LQTSNLNLGICIKGNCE 1287
: : : : :

```

```

Db      : : : : :
1613 ----LEKDEILKLYSCIIIGMISLPOESLAREASTGVLDKDMHGLITLMDSRIDLE 1667
: : : : :
Qy      : : : : :
1288 LPNKHFRSIRMNLNILE--AEHTDVYIAGLVLSKINRSNKRHNNMHFLLELILKTIQC 1346
: : : : :
Db      : : : : :
1668 EGQOIVRSVNLVAVKVEKSDQNTIISALLVLLQDSLLATASSPK--FSELVH----- 1718
: : : : :
Qy      : : : : :
1347 YQSHKALRDIIDSMIRIAPSLPD--LSINIVNPI-----ANGEPT----- 1388
: : : : :
Db      : : : : :
1719 ----KCLMRVRLPPTINSINLDRILDLHIFMKVFPREKJCKOCKSEPPRTLTKLH 1773
: : : : :
Qy      : : : : :
1389 NLCAIKILEVTEHHSSEITDAHDIY 1415
: : : : :
Db      : : : : :
1774 TLCKLK-----GPKIID-HLTM 1790
: : : : :

RESULT 12
US-10-408-765A-2422
: Sequence 2422, Application US/10408765A
: Publication No. US20040101874A1
: GENERAL INFORMATION:
: APPLICANT: Ghosh, Soumitra S.
: APPLICANT: Zhang, Bing
: APPLICANT: Fahy, Eoin D.
: APPLICANT: Gibson, Bradford W.
: APPLICANT: Taylor, Steven W.
: APPLICANT: Glenn, Gary M.
: APPLICANT: Warnock, Dale E.
: TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
: TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
: FILE REFERENCE: 660088 465
: CURRENT APPLICATION NUMBER: US/10/408, 765A
: CURRENT FILING DATE: 2003-04-04
: NUMBER OF SEQ ID NOS: 3077
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2422
: LENGTH: 306
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-408-765A-2422

Query Match 3.7%; Score 281.5; DB 4; Length 306;
Best Local Similarity 33.7%; Pred. No. 2,2e-10;
Matches 66; Conservative 46; Mismatches 81; Indels 5; Gaps 4;

Qy      : : : : :
1290 NKHRSIMRMLNILEAEHTDVYIAGLVLSKINRSNKRHNNMHFLLELILKTIQCYOH 1349
: : : : :
Db      : : : : :
109 DEHFKTILLLETLGDXEPTIRALAKVLRILRHQPAR--FQVAVELFVMTKLEAKHD 166
: : : : :
Qy      : : : : :
1350 -SKALRDISMIRIAPSLPDLISINIVNPIATGSEFPNLCIKILVTEHHSSEIT 1408
: : : : :
Db      : : : : :
167 PKREVNSAEBAASVLTSSIPBQCIKVLCPITQADYFNLAIKQTKIERVSKETL 226
: : : : :
Qy      : : : : :
1409 DAHLDIYFNARSADPTQSMVRKAAYFCIVKYFVGEKVKPKLSVLSKYRLNLNVY 1468
: : : : :
Db      : : : : :
227 NLLPEIMPGLIGQYDSESSVRKACVCLVAHNAVIGDE--LKHLSQLTOSKXKLNLXY 285
: : : : :
Qy      : : : : :
1469 IEKORNCISGGSGSTKNSAAS 1490
: : : : :
Db      : : : : :
286 I-KRAQTSGGADPTTVSGQS 306
: : : : :

RESULT 13
US-10-425-115-352121
: Sequence 352121, Application US/10425115
: Publication No. US20040214272A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Kovalic, David K.
: APPLICANT: Zhou, Yihua
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants
: FILE REFERENCE: 38-21(53222)B

```

```

; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 352121
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_8429C.1.pep
US-10-425-115-352121

```

Query Match 3.6%; Score 276; DB 4; Length 391;

Best Local Similarity 23.5%; Pred. No. 7.5e-10; Mismatches 145; Indels 66; Gaps 11;

Matches 91; Conservative 85; Mismatches 145; Indels 66; Gaps 11;

```

QY 42 IYCTDNGFLIDGMPMLTGSFHKIAQKSLKPSSELIKRLGSDPNATATVLPVHIDRLGD 101
DB 35 LTPAEVTAIVDTCTMDLTRDNFRIAGGLOALSVAAVAGDHFRIHNLVPAVERLGD 94
QY 102 SRDTRKQALLRDLMEHRLVLPQALIDKLATSCFHKNAKYREEFLOTIVNALHEYGT 161
DB 95 DKQPVHDAARQLLITLME--VSSPTTIVERAGSYATWTHKMRVRDEFRTVASAVGLFAS 152
QY 162 QQLSV-RVYIPVYCALLGDPVNVREAAIQTLVEIKYGVDRLRPDLRRMDVPASKLAM 220
DB 153 TEISLORVLLSPVLOLMDNSQVRDAALISCIEMVYKMGSOPEHELOR-HNLPSYMLKE 211
QY 221 LEQKPOVQOEGLLPSALKNTNGVGLDEADNIGLRERPTMRKRLPLHSAVSSLR-- 278
DB 212 INSRDLKIEPK-----VRSSDTTQYKAVESRGNAGMEPV 246
QY 279 ---PKPNVNDV---TGDAG-----AVTMSEFSSFEVVPQNLIFHAKMDDIY 320
DB 247 GSGTSGSVNDMENESSGSKRPLGRDSTKASRKAMSSSOSTVYLSRIH-----DIQ 301
QY 321 KOVLVIISDKNADWEKRVDAKKI--RALLILSYHTOPQVAVOL-----KELSLSF- 370
DB 302 IARLKSEKSELKQNIEMFMEKVELKLEFOSKTIEMQOTKLMEDRIKDEELNKMYA 361
QY 371 --VDIKELRSQVIREACTITAYMSK 395
DB 362 MDMDALPEELRAVYIARRKGLIDYFIK 388

```

RESULT 14
US-10-425-114-68783

Sequence 68783; Application US/10425114

Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jindong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E.

APPLICANT: Tabaka, Jack E.

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(5313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 68783

LENGTH: 407

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: 700352054_FLI.pep

US-10-425-114-68783

Query Match 3.6%; Score 276; DB 4; Length 407;

Best Local Similarity 23.5%; Pred. No. 7.9e-10; Mismatches 145; Indels 66; Gaps 11;

Matches 91; Conservative 85; Mismatches 145; Indels 66; Gaps 11;

```

QY 42 IYCTDNGFLIDGMPMLTGSFHKIAQKSLKPSSELIKRLGSDPNATATVLPVHIDRLGD 101
DB 51 LTPAEVTAIVDTCTMDLTRDNFRIAGGLOALSVAAVAGDHFRIHNLVPAVERLGD 110
QY 102 SRDTRKQALLRDLMEHRLVLPQALIDKLATSCFHKNAKYREEFLOTIVNALHEYGT 161
DB 111 DKQPVHDAARQLLITLME--VSSPTTIVERAGSYATWTHKMRVRDEFRTVASAVGLFAS 158
QY 162 QQLSV-RVYIPVYCALLGDPVNVREAAIQTLVEIKYGVDRLRPDLRRMDVPASKLAM 220
DB 169 TEISLORVLLSPVLOLMDNSQVRDAALISCIEMVYKMGSOPEHELOR-HNLPSYMLKE 227
QY 221 LEQKPOVQOEGLLPSALKNTNGVGLDEADNIGLRERPTMRKRLPLHSAVSSLR-- 278
DB 228 INSRDLKIEPK-----VRSSDTTQYKAVESRGNAGMEPV 262
QY 279 ---PKPNVNDV---TGDAG-----AVTMSEFSSFEVVPQNLIFHAKMDDIY 320
DB 263 GSGTSGSVNDMENESSGSKRPLGRDSTKASRKAMSSSOSTVYLSRIH-----DIQ 317
QY 321 KOVLVIISDKNADWEKRVDAKKI--RALLILSYHTOPQVAVOL-----KELSLSF- 370
DB 318 IARLKSEKSELKQNIEMFMEKVELKLEFOSKTIEMQOTKLMEDRIKDEELNKMYA 377
QY 371 --VDIKELRSQVIREACTITAYMSK 395
DB 378 MDMDALPEELRAVYIARRKGLIDYFIK 404

```

RESULT 15

US-10-437-963-123612

Sequence 123612; Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 123612

LENGTH: 1211

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_2642C.1.pep

US-10-437-963-123612

Query Match 3.5%; Score 264.5; DB 4; Length 1211;

Best Local Similarity 18.7%; Pred. No. 2.6e-08; Mismatches 260; Conservative 212; Mismatches 561; Indels 357; Gaps 54;

```

QY 171 PVYCALLGDPVNVREAAIQTLVEIKYGVDRLRPDLRRMDVPASKLAMLEQKPOVKQ 230
DB 79 PLVYGRGLGDDPDAVDAARFLVRLME----- 105
QY 231 EQLLPSALKNTN---GNGVGL---DEADNIGLRERPTMRKRLPLHSAVSSSLRKPKNV 283
DB 106 -----NKENNANMNGEKTTCISDVQDN-----RCMTIQIEPTDQAKSPKV 149
QY 284 NDVTGAGAVTMSEFSSFEVVPQNLIFHAKMDDIYKQVLVIISDKNADWEKRVDAKK 343
DB 150 KIVTRDASLLAAGAAYSAF---PML-----LQG 174
QY 344 IRALLILSYHTOPQVAVOLKELSLSFVDIKELRSQVIREACTITAYMSKTLRNKLD 403

```



```

Db      175 LVPLI-----TOLL--ORSSVVAQACHLNFSLKELLRDEP 211
Qy      404 FCMILLEHLNLIONSAKVIASTIALKYLITKTHAPKLKITYDTLNOGSKODISTL 463
Db      212 CAELLIFVLKNVITTHIVIAESSDNCIKEMLRCKARILPKIIEPAKDKSAVLARC 271
Qy      464 CELMVLLEEN-OTKALERNATVLRDLTKSIGA-----DCDARRHRYAWAFRRHF 516
Db      272 CEVALMELVDPIDRISVDLYEEFKCIEDATSEMINDEDAFTQORHL----- 323
Qy      517 PELADQYGTIDIAQALEREHREGGGGTGTGTGTAPEFRKRVISIGRPTGLQKTP 576
Db      324 -----SPVERKLMOPQSSGSCNST-----EIDKVK-----VD 351
Qy      577 SMRSISAVDTAAOARAKVAYOTLYSROKPLGPNNSNOASMTGAASGLPRPLNSNS 636
Db      352 SGTSFSSVDLOSVO-VKASVQHDMDTKIQPEKTD-----GTTTGS-----SFKDTS 399
Qy      637 GGTPTATPGSVTPPRGRAGV--SOSOPGSR--TSPTKLRDQYGGIGNYRGATGAI 691
Db      400 TLEKETSPDKGPDPCDGVNLSSCDPSASATQFLEPTSELPSDATV-----VTIV 452
Qy      692 PKKASGIPRTASRETSPTRSSGGLMKRSMYTGAGSRTPPERNNPVRPSAPARLLAQS 751
Db      453 QDKAECPNTEQLTOQVQVLGHPEDLNR-----SPPVSLKEG 491
Qy      752 REAHT-IGVDDDGQPPYVSGDYMRSGMRGRLMGRDESDDIDSEASVCSERSFDS 810
Db      492 NSLKNPFIQVNFDS-----SGKGAGQVKVHHTSAPN-----KSV 527
Qy      811 YTRGNKSNY-----SLSGHTRLDMSTORAPEDDIETIIQPCAS----- 849
Db      528 VRKEPRNNYIPNFRPPLSKQMTNMFYASNSDDLDKQLIGEMVSNMDVSSLTALSL 587
Qy      850 -----THMSERKDLISLTQYLA-----GKELTQQLKCVLDMFRKMFMDHTKYVSLF 900
Db      588 GLPKRSDMMKRYAFNFRSLRQSLLEQPRGIQEVANFQKVRFSRYLADPHHRIQAAL 647
Qy      901 DVTLELIVANETSRNGSSCLTRLFNKLTGTDLLNSHMSKIMKTLQVHEYPPTQLK 960
Db      648 SLSLEIMPVF-KKFEHYLDKTLPHVFSQ--NDPKESI KOQCSAILKLAGEIYISDLSL 705
Qy      961 ELPRISDSTQPTTKRIALRFLTLANT--YCKSDFPDQSQACERTVUKLAOLA 1017
Db      706 ALRTELEEQ--KSPKSKALIEF-----ANSSFVKCTVSSDNYSSSFLLKPMLEKALLF 758
Qy      1018 ADQKSMELRQARSCVALYNLNTPOMTLULADLPKYQDSARSCISHMRQSCNSG 1077
Db      759 KD-KNSKLKEVSAGFSSISYHDP--ASLISFLVLSMEQK-----RLRRAIKQLIPT 810
Qy      1078 ANSPSSSPLSSSPK--PLOSFSVG--PPAS--LOSHHQLSISTSPRSRQSEVQ 1128
Db      811 IESDLEFLQORRKHKPVPLFDSFVAGVOPYAKGLVIKONKHLHATYQ-----DCLQA 865
Qy      1129 ELFSSELDIQNIQKTSEIRHCFGGQYOTALAPNGFNHLOYHDQOOD--SCASLSS 1186
Db      866 DQVFDAL--HCLPRISLEVRERAGKIESE--SYDESNG--HDAEMMDKKSSDTRSS 917
Qy      1187 NSKTQSSANTQSTPESATRLDNLERERTONAKSPTDDAKIYVSIIMAEENGELILA 1246
Db      918 DPPRTFDYNIISDNTVESP--RKEGTDIKREESNEELN-----ITV-----RNRNVL 965
Qy      1247 SNLSESVVVALTLTDOPVELLOT--SLTNLIGICIGNCELPNKHFRSIRMRLNILE 1305
Db      966 -NSGQDGPBSISRTYHQDEMSQMEVEPEINGPSVSIK-GICKMLPYCIEIESLQNDLT 1023
Qy      1306 AEHTDVVIAGLHVLISKIMRSNKNHNMHLELILKXIOCYOHSKEALRD-IDSMIPRI 1364
Db      1024 IHRMSSSL--LEMDDDEESTR-----ELALTLLVEILEKORAMENCIDTLIVKL 1072
Qy      1365 APSLPLDLSINIVPV-----IATGEFTNLCAIKILLEVTE 1401

```

```

Db      1073 LHAATK-DAAKVVNQAHICLTTVYTOFDPULTCLRAIASQLANODEKYLIVINSLSKVI 1131
Qy      1402 HHSSEITDAHLDIVPNLRASADTQSMVKAAVFCIYVLYGEBKVRPKLSVINPSK 1461
Db      1132 RLSQDNLMHLSIFLPAALDAFENHSPYVRKAVVLCIVDTYKLGPAPFL-PYLERLDSAQ 1190
Qy      1462 VRLNLYTIEK 1471
Db      1191 LQLVTTYSR 1200

```

Search completed: November 23, 2005, 15:16:53
Job time : 212 secs